

RISET DAN PENGEMBANGAN DAKWAH ISLAMIAH ERA KECERDASAN BUATAN

Metodologi, Etika, Prototipe, Evaluasi Dampak,
dan Implementasi Nyata



Edisi 2026

KASMUI

RISET DAN PENGEMBANGAN DAKWAH ISLAMİYAH ERA KECERDASAN BUATAN

*Metodologi, Etika, Prototipe, Evaluasi Dampak, dan Implementasi
Nyata*

KASMUI

Edisi 2026

KATA PENGANTAR

Perubahan teknologi kecerdasan buatan tidak hanya menghadirkan perangkat baru, melainkan mengubah cara manusia mencari pengetahuan, membentuk kepercayaan, berkomunikasi, belajar, mengambil keputusan, dan membangun komunitas. Dakwah Islamiyah yang secara hakiki merupakan aktivitas mengajak kepada kebaikan, menyampaikan kebenaran dengan hikmah, serta membina manusia dan masyarakat, karena itu menghadapi medan baru yang jauh lebih kompleks daripada sekadar berpindah dari mimbar ke layar. Dalam ruang digital, pesan agama berhadapan dengan mesin pencari, sistem rekomendasi, model bahasa generatif, sintesis suara dan gambar, analitik perilaku, otomatisasi, serta ekonomi perhatian. Buku ini disusun untuk menempatkan semua perkembangan tersebut ke dalam kerangka riset dan pengembangan yang terukur, bertanggung jawab, dan tetap menegaskan otoritas manusia dalam urusan keagamaan.

Buku ini sengaja tidak dibangun dengan asumsi bahwa setiap teknologi baru pasti baik untuk dakwah. Teknologi adalah sarana yang nilai manfaatnya bergantung pada tujuan, desain, sumber data, tata kelola, konteks sosial, kompetensi pengguna, serta mekanisme pengawasan. UNESCO menekankan pendekatan berpusat pada manusia, martabat, transparansi, keadilan, perlindungan data, dan human oversight dalam pengembangan kecerdasan buatan. NIST menempatkan pengelolaan risiko sebagai proses sepanjang siklus hidup sistem. Dalam konteks dakwah, prinsip-prinsip tersebut menjadi semakin penting karena keluaran AI dapat menyentuh wilayah keyakinan, ibadah, hukum agama, moralitas, hubungan sosial, dan keputusan pribadi yang tidak layak diserahkan secara otomatis kepada sistem probabilistik.

Atas dasar itu, judul Riset dan Pengembangan Dakwah Islamiyah Era Kecerdasan Buatan dipilih untuk menegaskan dua kata kerja utama: meneliti dan mengembangkan. Meneliti berarti memulai dari masalah nyata, memetakan karakteristik mad'u, mengumpulkan data secara etis, menguji asumsi, membandingkan alternatif, mengukur perubahan, dan menyimpulkan berdasarkan bukti. Mengembangkan berarti merancang solusi yang dapat digunakan, misalnya korpus dakwah terverifikasi, mesin pencarian semantik, chatbot berbasis retrieval-augmented generation, sistem rekomendasi materi yang tidak manipulatif, transkripsi ceramah, alat bantu aksesibilitas, dashboard evaluasi program, serta protokol verifikasi konten. Kedua aktivitas harus berjalan bersama sehingga inovasi tidak berhenti pada demonstrasi teknologi, tetapi menghasilkan perbaikan praktik.

Pembaca akan menemukan bahwa model penelitian yang digunakan tidak dibatasi pada satu tradisi metodologis. Design Science Research berguna ketika sasaran penelitian adalah artefak digital; Educational Design Research relevan untuk sekolah, kampus, madrasah, dan program pembelajaran; participatory action

research penting ketika jamaah, dai, guru, mahasiswa, pengurus masjid, dan relawan harus menjadi mitra aktif; mixed methods diperlukan untuk menghubungkan data angka dan pengalaman; PRISMA membantu tinjauan sistematis; TAM dan UTAUT mendukung analisis penerimaan teknologi; sedangkan RE-AIM dan CFIR membantu menilai jangkauan, implementasi, adopsi, serta keberlanjutan program. Kombinasi ini memberi perangkat metodologis yang dapat dipilih sesuai pertanyaan riset, bukan dipaksakan sebagai resep tunggal.

Kebutuhan akan pendekatan ilmiah makin mendesak karena perkembangan AI berlangsung cepat. AI Index 2026 dari Stanford HAI menunjukkan percepatan kemampuan model sekaligus kesenjangan dalam pengukuran responsible AI. Pada saat yang sama, penggunaan AI untuk belajar makin luas sementara kebijakan dan literasi tidak selalu mengikuti kecepatannya. Kondisi tersebut menunjukkan dua tuntutan yang tampak berlawanan namun sesungguhnya saling melengkapi: lembaga dakwah perlu berinovasi agar tetap relevan dan melayani masyarakat, tetapi inovasi harus disertai pengendalian risiko, transparansi sumber, pembatasan fungsi, audit, dan pendidikan literasi AI.

Salah satu prinsip penting buku ini adalah pemisahan yang tegas antara alat bantu pengetahuan dan otoritas keagamaan. Model generatif dapat membantu mencari, merangkum, mengelompokkan, menerjemahkan, mentranskripsikan, atau menyusun alternatif presentasi, tetapi ia tidak memiliki tanggung jawab moral, sanad keilmuan, pemahaman konteks seperti manusia, ataupun otoritas normatif untuk menggantikan ulama, guru, dan lembaga yang berwenang. Karena itu, contoh chatbot keagamaan di buku ini selalu dirancang sebagai sistem grounded pada korpus terverifikasi, menampilkan sumber, membatasi ruang lingkup, mengakui ketidakpastian, dan mengalihkan pertanyaan sensitif atau kontekstual kepada manusia.

Buku ini juga berusaha menghadirkan contoh yang dapat benar-benar diterapkan. Di tingkat masjid, pembaca akan menemukan rancangan audit kebutuhan jamaah, pusat pengetahuan digital, layanan tanya-jawab non-fatwa, sistem dokumentasi khutbah dan kajian, serta evaluasi program. Di sekolah, terdapat desain microlearning, tutor pendamping guru, rubrik verifikasi materi, dan penelitian pretest-posttest. Di kampus, terdapat laboratorium dakwah digital, digital ethnography, eksperimen pesan secara etis, analisis jejaring, dan capstone project. Di tingkat organisasi, dibahas tata kelola portofolio AI, data governance, model card, incident log, audit konten, pusat kompetensi, dan roadmap kematangan organisasi.

Agar buku dapat dipakai sebagai buku ajar sekaligus manual kerja, setiap bab dilengkapi prosedur, tabel, gambar, contoh protokol, indikator, pertanyaan latihan, pembahasan, dan tugas pengembangan. Lampiran memuat instrumen yang dapat

diadaptasi: lembar analisis kebutuhan, informed consent, pedoman wawancara, kuesioner, rubrik fidelity, lembar uji RAG, daftar uji hallucination, risk register, data dictionary, model card ringkas, dan template laporan pilot. Pengguna tetap harus menyesuaikan instrumen dengan persetujuan etik, kebijakan lembaga, karakteristik populasi, dan regulasi yang berlaku.

Penulis menyadari bahwa kecerdasan buatan merupakan bidang yang berubah sangat cepat. Karena itu, buku ini lebih menekankan prinsip yang tahan perubahan: tujuan dakwah yang jelas, validitas sumber, metode ilmiah, perlindungan manusia, akuntabilitas, pengawasan manusia, keterlacakan, keterbukaan mengenai keterbatasan, evaluasi dampak, dan perbaikan berkelanjutan. Nama produk dan kemampuan model dapat berganti, tetapi prinsip-prinsip tersebut tetap menjadi fondasi untuk menilai apakah teknologi layak digunakan dalam pelayanan keagamaan.

Akhirnya, buku ini ditujukan bagi dosen, mahasiswa, dai, guru, pengurus masjid, peneliti komunikasi, pengembang perangkat lunak, pengelola lembaga dakwah, serta siapa pun yang ingin menghubungkan tradisi dakwah dengan metode riset dan rekayasa sistem yang bertanggung jawab. Harapannya, pembaca tidak hanya mampu menggunakan AI, tetapi mampu merumuskan masalah yang benar, memilih teknologi secara proporsional, membangun sistem yang dapat diuji, membaca data secara kritis, menjaga amanah pengetahuan, dan memastikan bahwa teknologi tetap menjadi alat yang membantu manusia menjalankan tanggung jawabnya.

Semarang, Agustus 2026

Penyusun,



KASMUI

DAFTAR ISI

KATA PENGANTAR	2
DAFTAR ISI	5
DAFTAR GAMBAR	17
DAFTAR TABEL	18
PETUNJUK PENGGUNAAN BUKU	21
LANDASAN NORMATIF RINGKAS	21
BAB 1 — FONDASI DAKWAH, RISET, DAN KECERDASAN BUATAN....	23
1.1 Dakwah sebagai Sistem Perubahan Sosial Berbasis Nilai	24
1.1.1 Landasan, konstruk, dan indikator	25
1.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	26
1.1.3 Evaluasi, risiko, dan keberlanjutan	27
1.2 Hikmah, Tabayyun, dan Amanah Informasi dalam Sistem Digital	28
1.2.1 Landasan, konstruk, dan indikator	29
1.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	30
1.2.3 Evaluasi, risiko, dan keberlanjutan	31
1.3 Dari Digitalisasi ke Transformasi Dakwah	32
1.3.1 Landasan, konstruk, dan indikator	33
1.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	34
1.3.3 Evaluasi, risiko, dan keberlanjutan	35
1.4 Apa yang Dapat dan Tidak Dapat Dilakukan AI	36
1.4.1 Landasan, konstruk, dan indikator	37
1.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	38
1.4.3 Evaluasi, risiko, dan keberlanjutan	39
1.5 Ekosistem AI: Data, Model, Platform, dan Manusia	40
1.5.1 Landasan, konstruk, dan indikator	41
1.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	42
1.5.3 Evaluasi, risiko, dan keberlanjutan	43
1.6 Kesenjangan Riset Dakwah AI	44
1.6.1 Landasan, konstruk, dan indikator	45
1.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	46
1.6.3 Evaluasi, risiko, dan keberlanjutan	47

1.7 Literasi AI untuk Dai dan Pengelola Dakwah.....	48
1.7.1 Landasan, konstruk, dan indikator	49
1.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	50
1.7.3 Evaluasi, risiko, dan keberlanjutan	51
1.8 Prinsip Human-in-the-Loop dan Otoritas Keagamaan.....	52
1.8.1 Landasan, konstruk, dan indikator	53
1.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	54
1.8.3 Evaluasi, risiko, dan keberlanjutan	55
LATIHAN DAN PEMBAHASAN BAB 1.....	56
SOAL PENGEMBANGAN BAB 1.....	63
BAB 2 — LANSKAP DIGITAL RELIGION DAN PERILAKU MADU	64
2.1 Digital Religion dan Kehidupan Beragama Hibrida	65
2.1.1 Landasan, konstruk, dan indikator	66
2.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	67
2.1.3 Evaluasi, risiko, dan keberlanjutan	68
2.2 Perubahan Otoritas Keagamaan di Ruang Algoritmik.....	69
2.2.1 Landasan, konstruk, dan indikator	70
2.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	71
2.2.3 Evaluasi, risiko, dan keberlanjutan	72
2.3 Segmentasi Madu Berbasis Kebutuhan, Bukan Manipulasi	73
2.3.1 Landasan, konstruk, dan indikator	74
2.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	75
2.3.3 Evaluasi, risiko, dan keberlanjutan	76
2.4 Ekonomi Perhatian, Algoritma Rekomendasi, dan Dakwah.....	77
2.4.1 Landasan, konstruk, dan indikator	78
2.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	79
2.4.3 Evaluasi, risiko, dan keberlanjutan	80
2.5 Disinformasi, Misinformasi, dan Klaim Agama Viral.....	81
2.5.1 Landasan, konstruk, dan indikator	82
2.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	83
2.5.3 Evaluasi, risiko, dan keberlanjutan	84
2.6 Bahasa Lokal, Inklusivitas, dan Aksesibilitas.....	85

2.6.1 Landasan, konstruk, dan indikator	86
2.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	87
2.6.3 Evaluasi, risiko, dan keberlanjutan	88
2.7 Digital Well-Being dan Batas Interaksi	89
2.7.1 Landasan, konstruk, dan indikator	90
2.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	91
2.7.3 Evaluasi, risiko, dan keberlanjutan	92
2.8 Metode Riset Perilaku Digital yang Etis	93
2.8.1 Landasan, konstruk, dan indikator	94
2.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	95
2.8.3 Evaluasi, risiko, dan keberlanjutan	96
LATIHAN DAN PEMBAHASAN BAB 2.....	97
SOAL PENGEMBANGAN BAB 2.....	104
BAB 3 — METODOLOGI RISET DAN PENGEMBANGAN DAKWAH AI	105
3.1 Merumuskan Masalah dan Theory of Change	106
3.1.1 Landasan, konstruk, dan indikator	107
3.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	108
3.1.3 Evaluasi, risiko, dan keberlanjutan	109
3.2 Design Science Research untuk Artefak Dakwah.....	110
3.2.1 Landasan, konstruk, dan indikator	111
3.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	112
3.2.3 Evaluasi, risiko, dan keberlanjutan	113
3.3 Educational Design Research untuk Pembelajaran Agama	114
3.3.1 Landasan, konstruk, dan indikator	115
3.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	116
3.3.3 Evaluasi, risiko, dan keberlanjutan	117
3.4 Participatory Action Research untuk Komunitas	118
3.4.1 Landasan, konstruk, dan indikator	119
3.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	120
3.4.3 Evaluasi, risiko, dan keberlanjutan	121
3.5 Mixed Methods dan Triangulasi	122
3.5.1 Landasan, konstruk, dan indikator	123

3.5.2	Prosedur riset dan pengembangan yang dapat diterapkan	124
3.5.3	Evaluasi, risiko, dan keberlanjutan	125
3.6	Systematic Review, Evidence Mapping, dan PRISMA	126
3.6.1	Landasan, konstruk, dan indikator	127
3.6.2	Prosedur riset dan pengembangan yang dapat diterapkan	128
3.6.3	Evaluasi, risiko, dan keberlanjutan	129
3.7	Instrumen, Validitas, Reliabilitas, dan Uji Coba.....	130
3.7.1	Landasan, konstruk, dan indikator	131
3.7.2	Prosedur riset dan pengembangan yang dapat diterapkan	132
3.7.3	Evaluasi, risiko, dan keberlanjutan	133
3.8	Analisis Data dan Pelaporan yang Dapat Direplikasi	134
3.8.1	Landasan, konstruk, dan indikator	135
3.8.2	Prosedur riset dan pengembangan yang dapat diterapkan	136
3.8.3	Evaluasi, risiko, dan keberlanjutan	137
	LATIHAN DAN PEMBAHASAN BAB 3.....	138
	SOAL PENGEMBANGAN BAB 3.....	144
	BAB 4 — DATA, KORPUS, DAN KNOWLEDGE ENGINEERING DAKWAH	146
4.1	Inventarisasi dan Kurasi Sumber Keagamaan.....	147
4.1.1	Landasan, konstruk, dan indikator	148
4.1.2	Prosedur riset dan pengembangan yang dapat diterapkan	149
4.1.3	Evaluasi, risiko, dan keberlanjutan	150
4.2	Taksonomi, Ontologi, dan Metadata Dakwah.....	151
4.2.1	Landasan, konstruk, dan indikator	152
4.2.2	Prosedur riset dan pengembangan yang dapat diterapkan	153
4.2.3	Evaluasi, risiko, dan keberlanjutan	154
4.3	Pembersihan, Normalisasi, dan Versioning Data.....	155
4.3.1	Landasan, konstruk, dan indikator	156
4.3.2	Prosedur riset dan pengembangan yang dapat diterapkan	157
4.3.3	Evaluasi, risiko, dan keberlanjutan	158
4.4	Embedding dan Pencarian Semantik.....	159
4.4.1	Landasan, konstruk, dan indikator	160

4.4.2	Prosedur riset dan pengembangan yang dapat diterapkan	161
4.4.3	Evaluasi, risiko, dan keberlanjutan	161
4.5	Retrieval-Augmented Generation untuk Tanya Jawab	162
4.5.1	Landasan, konstruk, dan indikator	164
4.5.2	Prosedur riset dan pengembangan yang dapat diterapkan	164
4.5.3	Evaluasi, risiko, dan keberlanjutan	165
4.6	Dataset Uji untuk Pertanyaan Keagamaan	166
4.6.1	Landasan, konstruk, dan indikator	167
4.6.2	Prosedur riset dan pengembangan yang dapat diterapkan	168
4.6.3	Evaluasi, risiko, dan keberlanjutan	169
4.7	Data Minimization dan Pelindungan Privasi.....	170
4.7.1	Landasan, konstruk, dan indikator	171
4.7.2	Prosedur riset dan pengembangan yang dapat diterapkan	172
4.7.3	Evaluasi, risiko, dan keberlanjutan	173
4.8	Knowledge Maintenance dan Pembaruan Berkala	174
4.8.1	Landasan, konstruk, dan indikator	175
4.8.2	Prosedur riset dan pengembangan yang dapat diterapkan	176
4.8.3	Evaluasi, risiko, dan keberlanjutan	177
	LATIHAN DAN PEMBAHASAN BAB 4.....	178
	SOAL PENGEMBANGAN BAB 4.....	185
BAB 5	— PENGEMBANGAN MEDIA, CHATBOT, DAN KONTEN GENERATIF.....	186
5.1	Prompt Engineering sebagai Spesifikasi Tugas	187
5.1.1	Landasan, konstruk, dan indikator	188
5.1.2	Prosedur riset dan pengembangan yang dapat diterapkan	189
5.1.3	Evaluasi, risiko, dan keberlanjutan	190
5.2	Chatbot Dakwah Non-Fatwa Berbasis RAG.....	191
5.2.1	Landasan, konstruk, dan indikator	192
5.2.2	Prosedur riset dan pengembangan yang dapat diterapkan	193
5.2.3	Evaluasi, risiko, dan keberlanjutan	194
5.3	Generasi Ringkasan, FAQ, dan Microlearning.....	195
5.3.1	Landasan, konstruk, dan indikator	196

5.3.2	Prosedur riset dan pengembangan yang dapat diterapkan	197
5.3.3	Evaluasi, risiko, dan keberlanjutan	198
5.4	Transkripsi, Caption, dan Terjemahan Ceramah.....	199
5.4.1	Landasan, konstruk, dan indikator	200
5.4.2	Prosedur riset dan pengembangan yang dapat diterapkan	201
5.4.3	Evaluasi, risiko, dan keberlanjutan	202
5.5	Visualisasi dan Desain Konten Berbantuan AI	203
5.5.1	Landasan, konstruk, dan indikator	204
5.5.2	Prosedur riset dan pengembangan yang dapat diterapkan	205
5.5.3	Evaluasi, risiko, dan keberlanjutan	206
5.6	Personalisasi yang Proporsional dan Transparan	207
5.6.1	Landasan, konstruk, dan indikator	208
5.6.2	Prosedur riset dan pengembangan yang dapat diterapkan	209
5.6.3	Evaluasi, risiko, dan keberlanjutan	209
5.7	Quality Assurance Konten AI	211
5.7.1	Landasan, konstruk, dan indikator	212
5.7.2	Prosedur riset dan pengembangan yang dapat diterapkan	212
5.7.3	Evaluasi, risiko, dan keberlanjutan	213
5.8	Publikasi, Provenance, dan Pelabelan Konten	214
5.8.1	Landasan, konstruk, dan indikator	215
5.8.2	Prosedur riset dan pengembangan yang dapat diterapkan	216
5.8.3	Evaluasi, risiko, dan keberlanjutan	217
LATIHAN DAN PEMBAHASAN BAB 5.....		218
SOAL PENGEMBANGAN BAB 5.....		225
BAB 6 — IMPLEMENTASI RIIL DI MASJID DAN KOMUNITAS.....		226
6.1	Audit Kebutuhan Jamaah 30 Hari	227
6.1.1	Landasan, konstruk, dan indikator	228
6.1.2	Prosedur riset dan pengembangan yang dapat diterapkan	229
6.1.3	Evaluasi, risiko, dan keberlanjutan	230
6.2	Pusat Pengetahuan Masjid Terverifikasi	231
6.2.1	Landasan, konstruk, dan indikator	232
6.2.2	Prosedur riset dan pengembangan yang dapat diterapkan	233

6.2.3 Evaluasi, risiko, dan keberlanjutan	234
6.3 Asisten Layanan Jamaah Non-Fatwa	235
6.3.1 Landasan, konstruk, dan indikator	236
6.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	237
6.3.3 Evaluasi, risiko, dan keberlanjutan	238
6.4 Arsip Khutbah dan Kajian yang Dapat Dicari	239
6.4.1 Landasan, konstruk, dan indikator	240
6.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	241
6.4.3 Evaluasi, risiko, dan keberlanjutan	241
6.5 Dakwah Berbasis Data Kebutuhan Sosial.....	242
6.5.1 Landasan, konstruk, dan indikator	244
6.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	244
6.5.3 Evaluasi, risiko, dan keberlanjutan	245
6.6 Program Literasi AI Jamaah dan Keluarga	246
6.6.1 Landasan, konstruk, dan indikator	247
6.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	248
6.6.3 Evaluasi, risiko, dan keberlanjutan	249
6.7 Volunteer Intelligence tanpa Surveillance	250
6.7.1 Landasan, konstruk, dan indikator	251
6.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	252
6.7.3 Evaluasi, risiko, dan keberlanjutan	253
6.8 Evaluasi Dampak Program Masjid.....	254
6.8.1 Landasan, konstruk, dan indikator	255
6.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	256
6.8.3 Evaluasi, risiko, dan keberlanjutan	257
LATIHAN DAN PEMBAHASAN BAB 6.....	258
SOAL PENGEMBANGAN BAB 6.....	265
BAB 7 — IMPLEMENTASI RIIL DI SEKOLAH DAN KAMPUS.....	266
7.1 AI Literacy dalam Pendidikan Agama	267
7.1.1 Landasan, konstruk, dan indikator	268
7.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	269
7.1.3 Evaluasi, risiko, dan keberlanjutan	270

7.2 Tutor Pendamping Guru Berbasis Korpus	271
7.2.1 Landasan, konstruk, dan indikator	272
7.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	273
7.2.3 Evaluasi, risiko, dan keberlanjutan	274
7.3 Microlearning dan Retrieval Practice.....	275
7.3.1 Landasan, konstruk, dan indikator	276
7.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	277
7.3.3 Evaluasi, risiko, dan keberlanjutan	278
7.4 Laboratorium Dakwah Digital Mahasiswa	279
7.4.1 Landasan, konstruk, dan indikator	280
7.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	281
7.4.3 Evaluasi, risiko, dan keberlanjutan	281
7.5 Digital Ethnography Komunitas Kampus	282
7.5.1 Landasan, konstruk, dan indikator	284
7.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	284
7.5.3 Evaluasi, risiko, dan keberlanjutan	285
7.6 Eksperimen Pesan Dakwah secara Etis.....	286
7.6.1 Landasan, konstruk, dan indikator	287
7.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	288
7.6.3 Evaluasi, risiko, dan keberlanjutan	289
7.7 Penilaian Autentik di Era GenAI	290
7.7.1 Landasan, konstruk, dan indikator	291
7.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	292
7.7.3 Evaluasi, risiko, dan keberlanjutan	293
7.8 Capstone Riset-Pengembangan Dakwah AI	294
7.8.1 Landasan, konstruk, dan indikator	295
7.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	296
7.8.3 Evaluasi, risiko, dan keberlanjutan	297
LATIHAN DAN PEMBAHASAN BAB 7.....	298
SOAL PENGEMBANGAN BAB 7.....	305
BAB 8 — IMPLEMENTASI RIIL PADA LEMBAGA DAKWAH	306
8.1 Strategi Portofolio AI Lembaga	307

8.1.1 Landasan, konstruk, dan indikator	308
8.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	309
8.1.3 Evaluasi, risiko, dan keberlanjutan	310
8.2 AI Center of Excellence Skala Lembaga	311
8.2.1 Landasan, konstruk, dan indikator	312
8.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	313
8.2.3 Evaluasi, risiko, dan keberlanjutan	314
8.3 SOP Pengadaan dan Evaluasi Vendor AI	315
8.3.1 Landasan, konstruk, dan indikator	316
8.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	317
8.3.3 Evaluasi, risiko, dan keberlanjutan	317
8.4 Dashboard Dakwah Berbasis Outcome.....	318
8.4.1 Landasan, konstruk, dan indikator	320
8.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	320
8.4.3 Evaluasi, risiko, dan keberlanjutan	321
8.5 Multibahasa dan Skala Distribusi Konten.....	322
8.5.1 Landasan, konstruk, dan indikator	323
8.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	324
8.5.3 Evaluasi, risiko, dan keberlanjutan	325
8.6 Knowledge Governance dan Editorial Board	326
8.6.1 Landasan, konstruk, dan indikator	327
8.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	328
8.6.3 Evaluasi, risiko, dan keberlanjutan	329
8.7 Incident Response untuk Kesalahan AI	330
8.7.1 Landasan, konstruk, dan indikator	331
8.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	332
8.7.3 Evaluasi, risiko, dan keberlanjutan	333
8.8 Model Kematangan Dakwah AI.....	334
8.8.1 Landasan, konstruk, dan indikator	335
8.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	336
8.8.3 Evaluasi, risiko, dan keberlanjutan	336
LATIHAN DAN PEMBAHASAN BAB 8.....	338

SOAL PENGEMBANGAN BAB 8.....	344
BAB 9 — ETIKA, HUKUM, KEAMANAN, DAN OTORITAS.....	345
9.1 Maqasid, Masalahah, dan Prinsip Do No Harm	346
9.1.1 Landasan, konstruk, dan indikator	347
9.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	348
9.1.3 Evaluasi, risiko, dan keberlanjutan	349
9.2 Human Dignity, Fairness, dan Non-Diskriminasi.....	350
9.2.1 Landasan, konstruk, dan indikator	351
9.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	352
9.2.3 Evaluasi, risiko, dan keberlanjutan	353
9.3 Pelindungan Data Pribadi dan Consent.....	354
9.3.1 Landasan, konstruk, dan indikator	355
9.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	356
9.3.3 Evaluasi, risiko, dan keberlanjutan	357
9.4 Transparansi, Explainability, dan Provenance	358
9.4.1 Landasan, konstruk, dan indikator	359
9.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	360
9.4.3 Evaluasi, risiko, dan keberlanjutan	360
9.5 Hallucination dan Reliability Engineering.....	361
9.5.1 Landasan, konstruk, dan indikator	363
9.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	363
9.5.3 Evaluasi, risiko, dan keberlanjutan	364
9.6 Deepfake, Sintesis Suara, dan Identitas Dai.....	365
9.6.1 Landasan, konstruk, dan indikator	366
9.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	367
9.6.3 Evaluasi, risiko, dan keberlanjutan	368
9.7 Keamanan Prompt, Data, dan Integrasi Sistem.....	369
9.7.1 Landasan, konstruk, dan indikator	370
9.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	371
9.7.3 Evaluasi, risiko, dan keberlanjutan	372
9.8 Batas Otoritas: AI Tidak Menggantikan Ulama.....	373
9.8.1 Landasan, konstruk, dan indikator	374

9.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	375
9.8.3 Evaluasi, risiko, dan keberlanjutan	376
LATIHAN DAN PEMBAHASAN BAB 9.....	377
SOAL PENGEMBANGAN BAB 9.....	384
BAB 10 — ROADMAP R&D, EVALUASI DAMPAK, DAN MASA DEPAN DAKWAH AI	385
10.1 Roadmap 90 Hari dari Masalah ke Pilot	386
10.1.1 Landasan, konstruk, dan indikator	387
10.1.2 Prosedur riset dan pengembangan yang dapat diterapkan	388
10.1.3 Evaluasi, risiko, dan keberlanjutan	389
10.2 Metrik Dampak: Reach, Learning, Trust, Action	390
10.2.1 Landasan, konstruk, dan indikator	391
10.2.2 Prosedur riset dan pengembangan yang dapat diterapkan	392
10.2.3 Evaluasi, risiko, dan keberlanjutan	393
10.3 TAM, UTAUT, dan Penerimaan Teknologi Dakwah.....	394
10.3.1 Landasan, konstruk, dan indikator	395
10.3.2 Prosedur riset dan pengembangan yang dapat diterapkan	396
10.3.3 Evaluasi, risiko, dan keberlanjutan	397
10.4 CFIR dan Faktor Implementasi.....	398
10.4.1 Landasan, konstruk, dan indikator	399
10.4.2 Prosedur riset dan pengembangan yang dapat diterapkan	400
10.4.3 Evaluasi, risiko, dan keberlanjutan	401
10.5 Replikasi, Skalabilitas, dan Keberlanjutan.....	402
10.5.1 Landasan, konstruk, dan indikator	403
10.5.2 Prosedur riset dan pengembangan yang dapat diterapkan	404
10.5.3 Evaluasi, risiko, dan keberlanjutan	405
10.6 Open Science dan Reproducible Da'wah Research.....	406
10.6.1 Landasan, konstruk, dan indikator	407
10.6.2 Prosedur riset dan pengembangan yang dapat diterapkan	408
10.6.3 Evaluasi, risiko, dan keberlanjutan	408
10.7 Skenario Masa Depan: Agentic AI dan Religious Information	409
10.7.1 Landasan, konstruk, dan indikator	411

10.7.2 Prosedur riset dan pengembangan yang dapat diterapkan	411
10.7.3 Evaluasi, risiko, dan keberlanjutan	412
10.8 Agenda Riset 2030 dan Prinsip Penutup	413
10.8.1 Landasan, konstruk, dan indikator	414
10.8.2 Prosedur riset dan pengembangan yang dapat diterapkan	415
10.8.3 Evaluasi, risiko, dan keberlanjutan	416
LATIHAN DAN PEMBAHASAN BAB 10.....	417
SOAL PENGEMBANGAN BAB 10.....	424
PORTOFOLIO 16 PROYEK RISET DAN PENGEMBANGAN YANG SIAP DITERAPKAN	425
DAFTAR PUSTAKA	428
GLOSARIUM	432
LAMPIRAN INSTRUMEN RISET DAN PENGEMBANGAN	437
A. Lembar Analisis Kebutuhan.....	437
B. Pedoman Wawancara Pengguna.....	438
C. Kuesioner Penerimaan Teknologi	439
D. Rubrik Fidelity Sumber.....	440
E. Test Set RAG.....	442
F. Checklist Hallucination.....	443
G. Risk Register	444
H. Data Dictionary	445
I. Model/System Card Ringkas	447
J. Laporan Pilot	448

DAFTAR GAMBAR

Gambar 1.1. Kerangka operasional Fondasi Dakwah, Riset, Dan Kecerdasan Buatan	20
Gambar 2.1. Kerangka operasional Lanskap Digital Religion Dan Perilaku Madu	59
Gambar 3.1. Kerangka operasional Metodologi Riset Dan Pengembangan Dakwah Ai ..	98
Gambar 4.1. Kerangka operasional Data, Korpus, Dan Knowledge Engineering Dakwah	136
Gambar 5.1. Kerangka operasional Pengembangan Media, Chatbot, Dan Konten Generatif	174
Gambar 6.1. Kerangka operasional Implementasi Riil Di Masjid Dan Komunitas.....	212
Gambar 7.1. Kerangka operasional Implementasi Riil Di Sekolah Dan Kampus	250
Gambar 8.1. Kerangka operasional Implementasi Riil Pada Lembaga Dakwah	288
Gambar 9.1. Kerangka operasional Etika, Hukum, Keamanan, Dan Otoritas.....	325
Gambar 10.1. Kerangka operasional Roadmap R&D, Evaluasi Dampak, Dan Masa Depan Dakwah Ai	362

DAFTAR TABEL

Tabel 1.1. Matriks kerja utama Bab 1	20
Tabel 1.2. Rencana tindakan riset: Dakwah sebagai Sistem Perubahan Sosial Berbasis Nilai	21
Tabel 1.3. Rencana tindakan riset: Hikmah, Tabayyun, dan Amanah Informasi dalam Sistem Digital.....	25
Tabel 1.4. Rencana tindakan riset: Dari Digitalisasi ke Transformasi Dakwah	29
Tabel 1.5. Rencana tindakan riset: Apa yang Dapat dan Tidak Dapat Dilakukan AI.....	33
Tabel 1.6. Rencana tindakan riset: Ekosistem AI: Data, Model, Platform, dan Manusia ..	36
Tabel 1.7. Rencana tindakan riset: Kesenjangan Riset Dakwah AI.....	40
Tabel 1.8. Rencana tindakan riset: Literasi AI untuk Dai dan Pengelola Dakwah	44
Tabel 1.9. Rencana tindakan riset: Prinsip Human-in-the-Loop dan Otoritas Keagamaan	47
Tabel 2.1. Matriks kerja utama Bab 2.....	59
Tabel 2.2. Rencana tindakan riset: Digital Religion dan Kehidupan Beragama Hibrida ..	60
Tabel 2.3. Rencana tindakan riset: Perubahan Otoritas Keagamaan di Ruang Algoritmik	64
Tabel 2.4. Rencana tindakan riset: Segmentasi Madu Berbasis Kebutuhan, Bukan Manipulasi	68
Tabel 2.5. Rencana tindakan riset: Ekonomi Perhatian, Algoritma Rekomendasi, dan Dakwah	71
Tabel 2.6. Rencana tindakan riset: Disinformasi, Misinformasi, dan Klaim Agama Viral	75
Tabel 2.7. Rencana tindakan riset: Bahasa Lokal, Inklusivitas, dan Aksesibilitas	79
Tabel 2.8. Rencana tindakan riset: Digital Well-Being dan Batas Interaksi.....	83
Tabel 2.9. Rencana tindakan riset: Metode Riset Perilaku Digital yang Etis	86
Tabel 3.1. Matriks kerja utama Bab 3.....	98
Tabel 3.2. Rencana tindakan riset: Merumuskan Masalah dan Theory of Change	99
Tabel 3.3. Rencana tindakan riset: Design Science Research untuk Artefak Dakwah	103
Tabel 3.4. Rencana tindakan riset: Educational Design Research untuk Pembelajaran Agama	107
Tabel 3.5. Rencana tindakan riset: Participatory Action Research untuk Komunitas	110
Tabel 3.6. Rencana tindakan riset: Mixed Methods dan Triangulasi.....	114
Tabel 3.7. Rencana tindakan riset: Systematic Review, Evidence Mapping, dan PRISMA	118
Tabel 3.8. Rencana tindakan riset: Instrumen, Validitas, Reliabilitas, dan Uji Coba	121
Tabel 3.9. Rencana tindakan riset: Analisis Data dan Pelaporan yang Dapat Direplikasi	125
Tabel 4.1. Matriks kerja utama Bab 4.....	136
Tabel 4.2. Rencana tindakan riset: Inventarisasi dan Kurasi Sumber Keagamaan	137
Tabel 4.3. Rencana tindakan riset: Taksonomi, Ontologi, dan Metadata Dakwah	141
Tabel 4.4. Rencana tindakan riset: Pembersihan, Normalisasi, dan Versioning Data.....	144
Tabel 4.5. Rencana tindakan riset: Embedding dan Pencarian Semantik	148
Tabel 4.6. Rencana tindakan riset: Retrieval-Augmented Generation untuk Tanya Jawab	152
Tabel 4.7. Rencana tindakan riset: Dataset Uji untuk Pertanyaan Keagamaan	155
Tabel 4.8. Rencana tindakan riset: Data Minimization dan Pelindungan Privasi	159
Tabel 4.9. Rencana tindakan riset: Knowledge Maintenance dan Pembaruan Berkala...	163
Tabel 5.1. Matriks kerja utama Bab 5.....	174

Tabel 5.2. Rencana tindakan riset: Prompt Engineering sebagai Spesifikasi Tugas	175
Tabel 5.3. Rencana tindakan riset: Chatbot Dakwah Non-Fatwa Berbasis RAG.....	179
Tabel 5.4. Rencana tindakan riset: Generasi Ringkasan, FAQ, dan Microlearning	182
Tabel 5.5. Rencana tindakan riset: Transkripsi, Caption, dan Terjemahan Ceramah.....	186
Tabel 5.6. Rencana tindakan riset: Visualisasi dan Desain Konten Berbantuan AI	190
Tabel 5.7. Rencana tindakan riset: Personalisasi yang Proporsional dan Transparan	193
Tabel 5.8. Rencana tindakan riset: Quality Assurance Konten AI.....	197
Tabel 5.9. Rencana tindakan riset: Publikasi, Provenance, dan Pelabelan Konten	201
Tabel 6.1. Matriks kerja utama Bab 6.....	212
Tabel 6.2. Rencana tindakan riset: Audit Kebutuhan Jamaah 30 Hari	213
Tabel 6.3. Rencana tindakan riset: Pusat Pengetahuan Masjid Terverifikasi	217
Tabel 6.4. Rencana tindakan riset: Asisten Layanan Jamaah Non-Fatwa	220
Tabel 6.5. Rencana tindakan riset: Arsip Khutbah dan Kajian yang Dapat Dicari.....	224
Tabel 6.6. Rencana tindakan riset: Dakwah Berbasis Data Kebutuhan Sosial	228
Tabel 6.7. Rencana tindakan riset: Program Literasi AI Jamaah dan Keluarga	231
Tabel 6.8. Rencana tindakan riset: Volunteer Intelligence tanpa Surveillance.....	235
Tabel 6.9. Rencana tindakan riset: Evaluasi Dampak Program Masjid	239
Tabel 7.1. Matriks kerja utama Bab 7.....	250
Tabel 7.2. Rencana tindakan riset: AI Literacy dalam Pendidikan Agama	251
Tabel 7.3. Rencana tindakan riset: Tutor Pendamping Guru Berbasis Korpus	255
Tabel 7.4. Rencana tindakan riset: Microlearning dan Retrieval Practice	258
Tabel 7.5. Rencana tindakan riset: Laboratorium Dakwah Digital Mahasiswa.....	262
Tabel 7.6. Rencana tindakan riset: Digital Ethnography Komunitas Kampus	266
Tabel 7.7. Rencana tindakan riset: Eksperimen Pesan Dakwah secara Etis	269
Tabel 7.8. Rencana tindakan riset: Penilaian Autentik di Era GenAI.....	273
Tabel 7.9. Rencana tindakan riset: Capstone Riset-Pengembangan Dakwah AI.....	276
Tabel 8.1. Matriks kerja utama Bab 8.....	288
Tabel 8.2. Rencana tindakan riset: Strategi Portofolio AI Lembaga	289
Tabel 8.3. Rencana tindakan riset: AI Center of Excellence Skala Lembaga.....	293
Tabel 8.4. Rencana tindakan riset: SOP Pengadaan dan Evaluasi Vendor AI.....	296
Tabel 8.5. Rencana tindakan riset: Dashboard Dakwah Berbasis Outcome	300
Tabel 8.6. Rencana tindakan riset: Multibahasa dan Skala Distribusi Konten	303
Tabel 8.7. Rencana tindakan riset: Knowledge Governance dan Editorial Board.....	307
Tabel 8.8. Rencana tindakan riset: Incident Response untuk Kesalahan AI.....	311
Tabel 8.9. Rencana tindakan riset: Model Kematangan Dakwah AI	314
Tabel 9.1. Matriks kerja utama Bab 9.....	325
Tabel 9.2. Rencana tindakan riset: Maqasid, Masalah, dan Prinsip Do No Harm.....	326
Tabel 9.3. Rencana tindakan riset: Human Dignity, Fairness, dan Non-Diskriminasi	330
Tabel 9.4. Rencana tindakan riset: Pelindungan Data Pribadi dan Consent	333
Tabel 9.5. Rencana tindakan riset: Transparansi, Explainability, dan Provenance	337
Tabel 9.6. Rencana tindakan riset: Hallucination dan Reliability Engineering	340
Tabel 9.7. Rencana tindakan riset: Deepfake, Sintesis Suara, dan Identitas Dai.....	344
Tabel 9.8. Rencana tindakan riset: Keamanan Prompt, Data, dan Integrasi Sistem.....	348

Tabel 9.9. Rencana tindakan riset: Batas Otoritas: AI Tidak Menggantikan Ulama.....	351
Tabel 10.1. Matriks kerja utama Bab 10.....	362
Tabel 10.2. Rencana tindakan riset: Roadmap 90 Hari dari Masalah ke Pilot.....	363
Tabel 10.3. Rencana tindakan riset: Metrik Dampak: Reach, Learning, Trust, Action...	367
Tabel 10.4. Rencana tindakan riset: TAM, UTAUT, dan Penerimaan Teknologi Dakwah	370
Tabel 10.5. Rencana tindakan riset: CFIR dan Faktor Implementasi	374
Tabel 10.6. Rencana tindakan riset: Replikasi, Skalabilitas, dan Keberlanjutan	378
Tabel 10.7. Rencana tindakan riset: Open Science dan Reproducible Da'wah Research	381
Tabel 10.8. Rencana tindakan riset: Skenario Masa Depan: Agentic AI dan Religious Information.....	385
Tabel 10.9. Rencana tindakan riset: Agenda Riset 2030 dan Prinsip Penutup	389
Tabel P.1. Portofolio proyek riset-pengembangan dakwah AI	400
Tabel Lampiran A. Template Lembar Analisis Kebutuhan	411
Tabel Lampiran B. Template Pedoman Wawancara Pengguna	412
Tabel Lampiran C. Template Kuesioner Penerimaan Teknologi.....	413
Tabel Lampiran D. Template Rubrik Fidelity Sumber	414
Tabel Lampiran E. Template Test Set RAG.....	415
Tabel Lampiran F. Template Checklist Hallucination	416
Tabel Lampiran G. Template Risk Register	416
Tabel Lampiran H. Template Data Dictionary	417
Tabel Lampiran I. Template Model/System Card Ringkas.....	418
Tabel Lampiran J. Template Laporan Pilot	419

PETUNJUK PENGGUNAAN BUKU

Buku ini dapat dibaca secara berurutan sebagai pengantar komprehensif atau digunakan secara modular. Bab 1-3 membangun dasar konseptual dan metodologis; Bab 4-5 membahas pembangunan pengetahuan dan artefak AI; Bab 6-8 menerjemahkan metode ke lingkungan masjid, sekolah/kampus, serta organisasi; Bab 9 memusatkan perhatian pada etika, hukum, keamanan, dan otoritas; Bab 10 menyatukan seluruh komponen ke dalam roadmap riset-pengembangan dan evaluasi dampak.

Untuk proyek nyata, pembaca dianjurkan memulai dari pertanyaan masalah, bukan dari nama aplikasi. Pilih satu persoalan terukur, tentukan populasi, tetapkan baseline, dokumentasikan sumber agama yang diizinkan, lakukan penilaian risiko, baru kemudian memilih teknologi. Seluruh contoh dalam buku harus dipahami sebagai rancangan penelitian yang perlu disesuaikan dengan persetujuan lembaga, kebijakan perlindungan data, dan otoritas keagamaan yang relevan.

Istilah AI dalam buku mencakup sistem prediktif, machine learning, natural language processing, model generatif, speech processing, computer vision, sistem rekomendasi, dan komponen retrieval. Tidak semua proyek memerlukan model besar. Dalam banyak kasus, pencarian semantik sederhana, basis pengetahuan terstruktur, transkripsi otomatis yang diperiksa manusia, atau dashboard analitik justru lebih aman, murah, dan tepat guna.

Tanda 'prototipe non-otoritatif' berarti sistem hanya membantu akses informasi dan tidak boleh diperlakukan sebagai pemberi fatwa, penentu hukum individual, atau pengganti konsultasi dengan manusia. Prinsip ini terutama berlaku untuk isu ibadah yang kompleks, pernikahan, perceraian, waris, konflik keluarga, kesehatan, keselamatan, hukum, dan keputusan lain yang memerlukan konteks personal serta pertanggungjawaban manusia.

LANDASAN NORMATIF RINGKAS

Landasan normatif berikut digunakan sebagai orientasi etis bagi proses riset dan pengembangan. Teks Arab diposisikan di kanan dan bukan dipakai untuk memberikan legitimasi otomatis kepada teknologi. Prinsip yang ditarik adalah hikmah, verifikasi, tanggung jawab atas pengetahuan, ucapan yang benar, dan ajakan kepada kebaikan.

أَدْعُ إِلَى سَبِيلِ رَبِّكَ بِالْحِكْمَةِ وَالْمَوْعِظَةِ الْحَسَنَةِ وَجَادِلْهُمْ بِالَّتِي هِيَ أَحْسَنُ

Serulah (manusia) ke jalan Tuhanmu dengan hikmah dan pengajaran yang baik, serta berdebatlah dengan mereka dengan cara yang lebih baik. (QS. An-Nahl [16]: 125)

يَا أَيُّهَا الَّذِينَ آمَنُوا إِن جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا أَن تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْحَبُوا عَلَى مَا فَعَلْتُمْ نَادِمِينَ

Wahai orang-orang yang beriman, jika seorang fasik datang kepadamu membawa berita penting, maka telitilah kebenarannya agar kamu tidak mencelakakan suatu kaum karena ketidaktahuan. (QS. Al-Hujurat [49]: 6)

وَلَا تَقْفُ مَا لَيْسَ لَكَ بِهِ عِلْمٌ إِنَّ السَّمْعَ وَالْبَصَرَ وَالْفُؤَادَ كُلُّ أُولَئِكَ كَانَ عَنْهُ مَسْئُولًا

Janganlah engkau mengikuti sesuatu yang tidak engkau ketahui. Sesungguhnya pendengaran, penglihatan, dan hati nurani, semua itu akan diminta pertanggungjawabannya. (QS. Al-Isra [17]: 36)

يَا أَيُّهَا الَّذِينَ آمَنُوا اتَّقُوا اللَّهَ وَقُولُوا قَوْلًا سَدِيدًا

Wahai orang-orang yang beriman, bertakwalah kepada Allah dan ucapkanlah perkataan yang benar. (QS. Al-Ahzab [33]: 70)

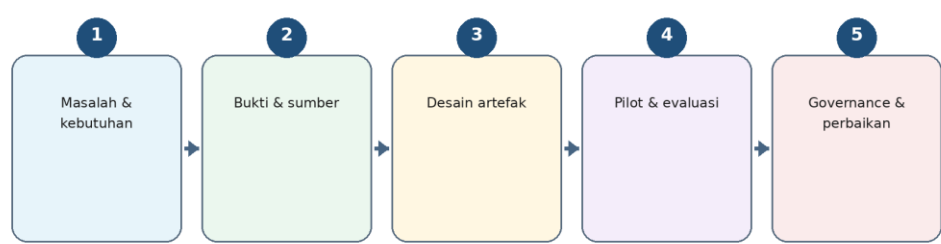
مَنْ دَعَا إِلَى هُدًى كَانَ لَهُ مِنَ الْأَجْرِ مِثْلُ أُجُورٍ مَنْ تَبِعَهُ، لَا يَنْقُصُ ذَلِكَ مِنْ أُجُورِهِمْ شَيْئًا

Siapa yang mengajak kepada petunjuk, ia memperoleh pahala seperti pahala orang-orang yang mengikutinya tanpa mengurangi pahala mereka sedikit pun. (HR. Muslim, no. 2674)

BAB 1 — FONDASI DAKWAH, RISET, DAN KECERDASAN BUATAN

Bab 1 membangun kompetensi untuk fondasi dakwah, riset, dan kecerdasan buatan. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 1



Gambar 1.1. Kerangka operasional Fondasi Dakwah, Riset, Dan Kecerdasan Buatan

Tabel 1.1. Matriks kerja utama Bab 1

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Desain	Bagaimana solusi bekerja dan di mana batasnya?	Spesifikasi + korpus	Prototipe	Uji internal
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

1.1 Dakwah sebagai Sistem Perubahan Sosial Berbasis Nilai

Dakwah sebagai Sistem Perubahan Sosial Berbasis Nilai ditempatkan sebagai unit kerja riset karena memahami dakwah sebagai proses komunikasi, pendidikan, pendampingan, pembinaan komunitas, dan perubahan perilaku yang harus dirumuskan dalam tujuan serta indikator. Fokus pada masjid, sekolah, kampus, dan organisasi mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Campbell & Bellar, 2023; Rogers, 2003. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi pemetaan masalah, wawancara, observasi, survei kebutuhan. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah logic model program dakwah, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.2. Rencana tindakan riset: Dakwah sebagai Sistem Perubahan Sosial Berbasis Nilai

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid, sekolah, kampus, dan organisasi; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi logic model program dakwah dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan pemetaan masalah, wawancara, observasi, survei kebutuhan.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: reduksi dakwah menjadi metrik popularitas; target yang tidak sensitif konteks.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam dakwah sebagai sistem perubahan sosial berbasis nilai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, sekolah, kampus, dan organisasi, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya

dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Campbell & Bellar, 2023; Rogers, 2003 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan pemetaan masalah, wawancara, observasi, survei kebutuhan. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi logic model program dakwah: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, sekolah, kampus, dan organisasi, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi reduksi dakwah menjadi metrik popularitas; target yang tidak sensitif konteks. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, logic model program dakwah memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, sekolah, kampus, dan organisasi, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Campbell & Bellar, 2023; Rogers, 2003).

Contoh penerapan riil: pada masjid, sekolah, kampus, dan organisasi, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas logic model program dakwah, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.2 Hikmah, Tabayyun, dan Amanah Informasi dalam Sistem Digital

Hikmah, Tabayyun, dan Amanah Informasi dalam Sistem Digital ditempatkan sebagai unit kerja riset karena menerjemahkan prinsip hikmah, verifikasi, tanggung jawab sumber, dan qaulan sadida ke prosedur pengembangan sistem. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021; Ji et al., 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi audit sumber, provenance log, review ahli. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah protokol verifikasi konten, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.3. Rencana tindakan riset: Hikmah, Tabayyun, dan Amanah Informasi dalam Sistem Digital

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi protokol verifikasi konten dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan audit sumber, provenance log, review ahli.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur akurasi sumber, keterlacakan, tingkat koreksi ahli.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: kutipan palsu, salah konteks, otoritas semu.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam hikmah, tabayyun, dan amanah informasi dalam sistem digital, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah akurasi sumber, keterlacakan, tingkat koreksi ahli. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus

dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021; Ji et al., 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan audit sumber, provenance log, review ahli. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi protokol verifikasi konten: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator akurasi sumber, keterlacakan, tingkat koreksi ahli dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa akurasi sumber, keterlacakan, tingkat koreksi ahli; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi kutipan palsu, salah konteks, otoritas semu. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, protokol verifikasi konten memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021; Ji et al., 2023).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas protokol verifikasi konten, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.3 Dari Digitalisasi ke Transformasi Dakwah

Dari Digitalisasi ke Transformasi Dakwah ditempatkan sebagai unit kerja riset karena membedakan unggah konten, otomasi proses, dan transformasi layanan berbasis data. Fokus pada lembaga dakwah mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Campbell & Tsuria, 2022; Bunt, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi audit proses dan pain-point. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah peta transformasi layanan, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.4. Rencana tindakan riset: Dari Digitalisasi ke Transformasi Dakwah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei	Ringkasan	Masalah spesifik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	pada lembaga dakwah; prioritaskan satu masalah.	baseline, consent, catatan sampling	dan terukur
2. Desain	Susun spesifikasi peta transformasi layanan dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan audit proses dan pain-point.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur waktu layanan, aksesibilitas, kepuasan, kualitas.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: otomasi proses buruk; ketergantungan vendor.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam dari digitalisasi ke transformasi dakwah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah waktu layanan, aksesibilitas, kepuasan, kualitas. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga dakwah, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi

sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Campbell & Tsuria, 2022; Bunt, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan audit proses dan pain-point. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi peta transformasi layanan: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga dakwah, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator waktu layanan, aksesibilitas, kepuasan, kualitas dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem

'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa waktu layanan, aksesibilitas, kepuasan, kualitas; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi otomasi proses buruk; ketergantungan vendor. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, peta transformasi layanan memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga dakwah, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Campbell & Tsuria, 2022; Bunt, 2018).

Contoh penerapan riil: pada lembaga dakwah, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas peta transformasi layanan, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.4 Apa yang Dapat dan Tidak Dapat Dilakukan AI

Apa yang Dapat dan Tidak Dapat Dilakukan AI ditempatkan sebagai unit kerja riset karena menguraikan kemampuan prediksi, klasifikasi, generasi, retrieval, transkripsi, terjemahan, dan keterbatasan epistemik model. Fokus pada masjid dan lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Vaswani et al., 2017; Lewis et al., 2020; NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi uji tugas, benchmark lokal, expert review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah matriks kesesuaian tugas-AI, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.5. Rencana tindakan riset: Apa yang Dapat dan Tidak Dapat Dilakukan AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid dan lembaga;	Ringkasan baseline,	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	prioritaskan satu masalah.	consent, catatan sampling	
2. Desain	Susun spesifikasi matriks kesesuaian tugas-AI dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan uji tugas, benchmark lokal, expert review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur akurasi, coverage, abstention rate.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: hallucination, bias, overtrust.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam apa yang dapat dan tidak dapat dilakukan ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah akurasi, coverage, abstention rate. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid dan lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi

sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Vaswani et al., 2017; Lewis et al., 2020; NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan uji tugas, benchmark lokal, expert review. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi matriks kesesuaian tugas-AI: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid dan lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator akurasi, coverage, abstention rate dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem

'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa akurasi, coverage, abstention rate; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi hallucination, bias, overtrust. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, matriks kesesuaian tugas-AI memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid dan lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar

untuk mengambil keputusan tersebut (Vaswani et al., 2017; Lewis et al., 2020; NIST, 2024).

Contoh penerapan riil: pada masjid dan lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas matriks kesesuaian tugas-AI, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.5 Ekosistem AI: Data, Model, Platform, dan Manusia

Ekosistem AI: Data, Model, Platform, dan Manusia ditempatkan sebagai unit kerja riset karena memahami AI sebagai sistem sosioteknis yang melibatkan data, proses, infrastruktur, pengguna, pengelola, dan tata kelola. Fokus pada organisasi mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain ISO/IEC 42001, 2023; NIST, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi stakeholder mapping dan data flow mapping. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah arsitektur sosioteknis dakwah AI, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.6. Rencana tindakan riset: Ekosistem AI: Data, Model, Platform, dan Manusia

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada organisasi; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		sampling	
2. Desain	Susun spesifikasi arsitektur sosioteknis dakwah AI dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan stakeholder mapping dan data flow mapping.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur keterlacakan data, peran jelas, SLA.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: shadow AI, akses berlebih, data leakage.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam ekosistem ai: data, model, platform, dan manusia, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah keterlacakan data, peran jelas, SLA. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada organisasi, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. ISO/IEC 42001, 2023; NIST, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan stakeholder mapping dan data flow mapping. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi arsitektur sosioteknis dakwah AI: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada organisasi, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator keterlacakan data, peran jelas, SLA dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa keterlacakan data, peran jelas, SLA; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi shadow AI, akses berlebih, data leakage. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, arsitektur sosioteknis dakwah AI memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks organisasi, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (ISO/IEC 42001, 2023; NIST, 2023).

Contoh penerapan riil: pada organisasi, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas arsitektur sosioteknis dakwah AI, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.6 Kesenjangan Riset Dakwah AI

Kesenjangan Riset Dakwah AI ditempatkan sebagai unit kerja riset karena mengidentifikasi area yang banyak dibahas secara normatif tetapi belum diuji secara empiris, terutama dampak, reliabilitas, dan implementasi jangka panjang. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Page et al., 2021; Winter, 2026. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi systematic mapping review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah agenda riset prioritas, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.7. Rencana tindakan riset: Kesenjangan Riset Dakwah AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi agenda riset prioritas dan batas	Spesifikasi, daftar sumber,	Disetujui tim dan substansi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	fungsi.	risk register	teknis
3. Prototipe	Bangun versi minimum menggunakan systematic mapping review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur jumlah pertanyaan teruji, replikasi, kualitas bukti.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: trend chasing dan publication bias.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam kesenjangan riset dakwah ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah jumlah pertanyaan teruji, replikasi, kualitas bukti. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Page et al., 2021; Winter, 2026 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan systematic mapping review. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi agenda riset prioritas: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator jumlah pertanyaan teruji, replikasi, kualitas bukti dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena

R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa jumlah pertanyaan teruji, replikasi, kualitas bukti; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi trend chasing dan publication bias. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, agenda riset prioritas memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Page et al., 2021; Winter, 2026).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-

keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas agenda riset prioritas, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.7 Literasi AI untuk Dai dan Pengelola Dakwah

Literasi AI untuk Dai dan Pengelola Dakwah ditempatkan sebagai unit kerja riset karena membangun kompetensi memahami cara kerja, batasan, prompting, verifikasi, privasi, dan pelaporan insiden. Fokus pada masjid dan lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Miao & Holmes, 2023; UNESCO, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi pre-post assessment dan observasi kinerja. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah kurikulum literasi AI dakwah, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.8. Rencana tindakan riset: Literasi AI untuk Dai dan Pengelola Dakwah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid dan lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi kurikulum literasi AI dakwah dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan pre-post	Version log, test set, defect	Lulus uji internal

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	assessment dan observasi kinerja.	log	minimum
4. Pilot	Uji terbatas; ukur gain score, error detection, compliance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: automation bias dan deskilling.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam literasi ai untuk dai dan pengelola dakwah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah gain score, error detection, compliance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid dan lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Miao & Holmes, 2023; UNESCO, 2024 memberi kerangka awal, sementara

konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan pre-post assessment dan observasi kinerja. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi kurikulum literasi AI dakwah: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid dan lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator gain score, error detection, compliance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa gain score, error detection, compliance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi automation bias dan deskilling. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, kurikulum literasi AI dakwah memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid dan lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Miao & Holmes, 2023; UNESCO, 2024).

Contoh penerapan riil: pada masjid dan lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas kurikulum literasi AI dakwah,

dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

1.8 Prinsip Human-in-the-Loop dan Otoritas Keagamaan

Prinsip Human-in-the-Loop dan Otoritas Keagamaan ditempatkan sebagai unit kerja riset karena menetapkan titik keputusan yang wajib dilakukan manusia dan proses eskalasi untuk pertanyaan yang memerlukan otoritas. Fokus pada lembaga fatwa/non-fatwa mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Bunt, 2018; Whyte, 2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi workflow analysis dan expert consensus. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah peta eskalasi manusia-AI, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 1.9. Rencana tindakan riset: Prinsip Human-in-the-Loop dan Otoritas Keagamaan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga fatwa/non-fatwa; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi peta eskalasi manusia-AI dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan workflow analysis dan expert consensus.	Version log, test set, defect log	Lulus uji internal minimum

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
4. Pilot	Uji terbatas; ukur abstention benar, waktu eskalasi, auditability.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: AI dianggap mufti; keputusan tanpa konteks.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

1.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam prinsip human-in-the-loop dan otoritas keagamaan, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah abstention benar, waktu eskalasi, auditability. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga fatwa/non-fatwa, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Bunt, 2018; Whyte, 2022 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan

kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

1.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan workflow analysis dan expert consensus. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi peta eskalasi manusia-AI: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga fatwa/non-fatwa, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator abstention benar, waktu eskalasi, auditability dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

1.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa abstention benar, waktu eskalasi, auditability; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi AI dianggap mufti; keputusan tanpa konteks. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, peta eskalasi manusia-AI memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga fatwa/non-fatwa, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Bunt, 2018; Whyte, 2022).

Contoh penerapan riil: pada lembaga fatwa/non-fatwa, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas peta eskalasi manusia-AI,

dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 1

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, sekolah, kampus, dan organisasi, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami dakwah sebagai proses komunikasi, pendidikan, pendampingan, pembinaan komunitas, dan perubahan perilaku yang harus dirumuskan dalam tujuan serta indikator' sebagai contoh. Bukti dapat berasal dari pemetaan masalah, wawancara, observasi, survei kebutuhan, sedangkan produk atau intervensinya dapat berbentuk logic model program dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti reduksi dakwah menjadi metrik popularitas; target yang tidak sensitif konteks perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerjemahkan prinsip hikmah, verifikasi, tanggung jawab sumber, dan qaulan sadida ke prosedur pengembangan sistem' sebagai contoh. Bukti dapat berasal dari audit sumber, provenance log, review ahli, sedangkan produk atau intervensinya dapat berbentuk protokol verifikasi konten. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti akurasi sumber, keterlacakan, tingkat koreksi ahli harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kutipan palsu, salah konteks, otoritas semu perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga dakwah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membedakan unggah konten, otomatisasi proses, dan transformasi layanan berbasis data' sebagai contoh. Bukti dapat berasal dari audit proses dan pain-point, sedangkan produk atau intervensinya dapat berbentuk peta transformasi layanan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti waktu layanan, aksesibilitas, kepuasan, kualitas harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti otomatisasi proses buruk; ketergantungan vendor perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid dan lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguraikan kemampuan prediksi, klasifikasi, generasi, retrieval, transkripsi, terjemahan, dan keterbatasan epistemik model' sebagai contoh. Bukti dapat berasal dari uji tugas, benchmark lokal, expert review, sedangkan produk atau intervensinya dapat berbentuk matriks kesesuaian tugas-AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti akurasi, coverage, abstention rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti hallucination, bias, overtrust perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks organisasi, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami AI sebagai sistem sosioteknis yang melibatkan data, proses, infrastruktur, pengguna, pengelola, dan tata kelola' sebagai contoh. Bukti dapat berasal dari stakeholder mapping dan data flow mapping, sedangkan produk atau intervensinya dapat berbentuk arsitektur sosioteknis dakwah AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti keterlacakan data, peran jelas, SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti shadow AI, akses berlebih, data leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengidentifikasi area yang banyak dibahas secara normatif tetapi belum diuji secara empiris, terutama dampak, reliabilitas, dan implementasi jangka panjang' sebagai contoh. Bukti dapat berasal dari systematic mapping review, sedangkan produk atau intervensinya dapat berbentuk agenda riset prioritas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti jumlah pertanyaan teruji, replikasi, kualitas bukti harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti trend chasing dan publication bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid dan lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun kompetensi memahami cara kerja, batasan, prompting, verifikasi, privasi, dan pelaporan insiden' sebagai contoh. Bukti dapat berasal dari pre-post assessment dan observasi kinerja, sedangkan produk atau intervensinya dapat berbentuk kurikulum literasi AI dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti gain score, error detection, compliance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti automation bias dan deskilling perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga fatwa/non-fatwa, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menetapkan titik keputusan yang wajib dilakukan manusia dan proses eskalasi untuk pertanyaan yang memerlukan otoritas' sebagai contoh. Bukti dapat berasal dari workflow analysis dan expert consensus, sedangkan produk atau intervensinya dapat berbentuk peta eskalasi manusia-AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti abstention benar, waktu eskalasi, auditability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI dianggap mufti; keputusan tanpa konteks perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, sekolah, kampus, dan organisasi, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami dakwah sebagai proses komunikasi, pendidikan, pendampingan, pembinaan komunitas, dan perubahan perilaku yang harus dirumuskan dalam tujuan serta indikator' sebagai contoh. Bukti dapat berasal dari pemetaan masalah, wawancara, observasi, survei kebutuhan, sedangkan produk atau intervensinya dapat berbentuk logic model program dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti reduksi dakwah menjadi metrik popularitas; target yang tidak sensitif konteks perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerjemahkan prinsip hikmah, verifikasi, tanggung jawab sumber, dan qaulan sadida ke prosedur pengembangan sistem' sebagai contoh. Bukti dapat berasal dari audit sumber, provenance log, review ahli, sedangkan produk atau intervensinya dapat berbentuk protokol verifikasi konten. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti akurasi sumber, keterlacakan, tingkat koreksi ahli harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kutipan palsu, salah konteks, otoritas semu perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga dakwah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membedakan unggah konten, otomasi proses, dan transformasi layanan berbasis data' sebagai contoh. Bukti dapat berasal dari audit proses dan pain-point, sedangkan produk atau intervensinya dapat berbentuk peta transformasi layanan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti waktu layanan, aksesibilitas, kepuasan, kualitas harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti otomasi proses buruk; ketergantungan vendor perlu ditulis sebagai skenario konkret, diberi

pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid dan lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguraikan kemampuan prediksi, klasifikasi, generasi, retrieval, transkripsi, terjemahan, dan keterbatasan epistemik model' sebagai contoh. Bukti dapat berasal dari uji tugas, benchmark lokal, expert review, sedangkan produk atau intervensinya dapat berbentuk matriks kesesuaian tugas-AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti akurasi, coverage, abstention rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti hallucination, bias, overtrust perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks organisasi, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami AI sebagai sistem sosioteknis yang melibatkan data, proses, infrastruktur, pengguna, pengelola, dan tata kelola' sebagai contoh. Bukti dapat berasal dari stakeholder mapping dan data flow mapping, sedangkan produk atau intervensinya dapat berbentuk arsitektur sosioteknis dakwah AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti keterlacakan data, peran jelas, SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti shadow AI, akses berlebih, data leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengidentifikasi area yang banyak dibahas secara normatif tetapi belum diuji secara empiris, terutama dampak, reliabilitas, dan implementasi jangka panjang' sebagai contoh. Bukti dapat berasal dari systematic mapping review, sedangkan produk atau intervensinya dapat berbentuk agenda riset prioritas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti jumlah pertanyaan teruji, replikasi, kualitas bukti harus dibaca bersama

temuan kualitatif dan catatan insiden. Risiko seperti trend chasing dan publication bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid dan lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun kompetensi memahami cara kerja, batasan, prompting, verifikasi, privasi, dan pelaporan insiden' sebagai contoh. Bukti dapat berasal dari pre-post assessment dan observasi kinerja, sedangkan produk atau intervensinya dapat berbentuk kurikulum literasi AI dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti gain score, error detection, compliance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti automation bias dan deskilling perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga fatwa/non-fatwa, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menetapkan titik keputusan yang wajib dilakukan manusia dan proses eskalasi untuk pertanyaan yang memerlukan otoritas' sebagai contoh. Bukti dapat berasal dari workflow analysis dan expert consensus, sedangkan produk atau intervensinya dapat berbentuk peta eskalasi manusia-AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti abstention benar, waktu eskalasi, auditability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI dianggap mufti; keputusan tanpa konteks perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, sekolah, kampus, dan organisasi, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami dakwah sebagai proses komunikasi, pendidikan, pendampingan, pembinaan komunitas, dan perubahan perilaku yang harus dirumuskan dalam tujuan serta indikator' sebagai contoh. Bukti dapat berasal dari pemetaan masalah, wawancara, observasi, survei

kebutuhan, sedangkan produk atau intervensinya dapat berbentuk logic model program dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti jangkauan, pemahaman, keterlibatan, perubahan perilaku prososial harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti reduksi dakwah menjadi metrik popularitas; target yang tidak sensitif konteks perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerjemahkan prinsip hikmah, verifikasi, tanggung jawab sumber, dan qaulan sadida ke prosedur pengembangan sistem' sebagai contoh. Bukti dapat berasal dari audit sumber, provenance log, review ahli, sedangkan produk atau intervensinya dapat berbentuk protokol verifikasi konten. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti akurasi sumber, keterlacakan, tingkat koreksi ahli harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kutipan palsu, salah konteks, otoritas semu perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga dakwah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membedakan unggah konten, otomasi proses, dan transformasi layanan berbasis data' sebagai contoh. Bukti dapat berasal dari audit proses dan pain-point, sedangkan produk atau intervensinya dapat berbentuk peta transformasi layanan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti waktu layanan, aksesibilitas, kepuasan, kualitas harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti otomasi proses buruk; ketergantungan vendor perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid dan lembaga, tujuan yang terukur, data yang proporsional, dan

kontrol risiko. Gunakan fokus 'menguraikan kemampuan prediksi, klasifikasi, generasi, retrieval, transkripsi, terjemahan, dan keterbatasan epistemik model' sebagai contoh. Bukti dapat berasal dari uji tugas, benchmark lokal, expert review, sedangkan produk atau intervensinya dapat berbentuk matriks kesesuaian tugas-AI. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti akurasi, coverage, abstention rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti hallucination, bias, overtrust perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

SOAL PENGEMBANGAN BAB 1

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 1 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 1, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 2 — LANSKAP DIGITAL RELIGION DAN PERILAKU MADU

Bab 2 membangun kompetensi untuk lanskap digital religion dan perilaku madu. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 2



Gambar 2.1. Kerangka operasional Lanskap Digital Religion Dan Perilaku Madu

Tabel 2.1. Matriks kerja utama Bab 2

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Desain	Bagaimana solusi bekerja dan di mana batasnya?	Spesifikasi + korpus	Prototipe	Uji internal
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

2.1 Digital Religion dan Kehidupan Beragama Hibrida

Digital Religion dan Kehidupan Beragama Hibrida ditempatkan sebagai unit kerja riset karena menganalisis hubungan praktik online dan offline serta perubahan komunitas, identitas, autentisitas, dan otoritas. Fokus pada komunitas mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Campbell & Bellar, 2023; Campbell & Tsuria, 2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi digital ethnography, wawancara, diary study. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah peta perjalanan pengalaman beragama hibrida, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.2. Rencana tindakan riset: Digital Religion dan Kehidupan Beragama Hibrida

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada komunitas; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi peta perjalanan pengalaman beragama hibrida dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan digital ethnography, wawancara, diary study.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur keterlibatan lintas kanal, kualitas relasi.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: reduksi pengalaman ke klik; pengamatan invasif.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam digital religion dan kehidupan beragama hibrida, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah keterlibatan lintas kanal, kualitas relasi. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada komunitas, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh

mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Campbell & Bellar, 2023; Campbell & Tsuria, 2022 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan digital ethnography, wawancara, diary study. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi peta perjalanan pengalaman beragama hibrida: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada komunitas, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator keterlibatan lintas kanal, kualitas relasi dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa keterlibatan lintas kanal, kualitas relasi; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi reduksi pengalaman ke klik; pengamatan invasif. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, peta perjalanan pengalaman beragama hibrida memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks komunitas, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Campbell & Bellar, 2023; Campbell & Tsuria, 2022).

Contoh penerapan riil: pada komunitas, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas peta perjalanan pengalaman beragama hibrida, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.2 Perubahan Otoritas Keagamaan di Ruang Algoritmik

Perubahan Otoritas Keagamaan di Ruang Algoritmik ditempatkan sebagai unit kerja riset karena menelaah bagaimana visibilitas, reputasi, algoritma, dan jaringan memengaruhi siapa yang dianggap otoritatif. Fokus pada media sosial mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Bunt, 2018; Whyte, 2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi network analysis, content analysis, wawancara. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah peta otoritas dan mekanisme kredibilitas, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.3. Rencana tindakan riset: Perubahan Otoritas Keagamaan di Ruang Algoritmik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada media sosial; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi peta otoritas dan mekanisme kredibilitas dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan network analysis, content analysis, wawancara.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur credential awareness, source checking.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: pseudo-clerics, fatwa shopping, authority by virality.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam perubahan otoritas keagamaan di ruang algoritmik, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah credential awareness, source checking. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada media sosial, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi

analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Bunt, 2018; Whyte, 2022 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan network analysis, content analysis, wawancara. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi peta otoritas dan mekanisme kredibilitas: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada media sosial, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator credential awareness, source checking dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa credential awareness, source checking; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi pseudo-clerics, fatwa shopping, authority by virality. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, peta otoritas dan mekanisme kredibilitas memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks media sosial, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Bunt, 2018; Whyte, 2022).

Contoh penerapan riil: pada media sosial, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas peta otoritas dan mekanisme kredibilitas, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.3 Segmentasi Madu Berbasis Kebutuhan, Bukan Manipulasi

Segmentasi Madu Berbasis Kebutuhan, Bukan Manipulasi ditempatkan sebagai unit kerja riset karena mengelompokkan kebutuhan audiens secara etis untuk penyesuaian materi tanpa mengeksploitasi kerentanan. Fokus pada masjid, sekolah, kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021; OECD, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi survei, clustering anonim, persona research. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah persona kebutuhan dakwah, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.4. Rencana tindakan riset: Segmentasi Madu Berbasis Kebutuhan, Bukan Manipulasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid, sekolah, kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi persona kebutuhan dakwah dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan survei, clustering anonim, persona research.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur relevansi materi, akses, kepuasan.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: profiling sensitif, discrimination.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam segmentasi madu berbasis kebutuhan, bukan manipulasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah relevansi materi, akses, kepuasan. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, sekolah, kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus

dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021; OECD, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan survei, clustering anonim, persona research. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi persona kebutuhan dakwah: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, sekolah, kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator relevansi materi, akses, kepuasan dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa relevansi materi, akses, kepuasan; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi profiling sensitif, discrimination. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, persona kebutuhan dakwah memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, sekolah, kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021; OECD, 2024).

Contoh penerapan riil: pada masjid, sekolah, kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas persona kebutuhan dakwah, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.4 Ekonomi Perhatian, Algoritma Rekomendasi, dan Dakwah

Ekonomi Perhatian, Algoritma Rekomendasi, dan Dakwah ditempatkan sebagai unit kerja riset karena memahami dampak ranking, engagement, dan insentif platform terhadap bentuk pesan keagamaan. Fokus pada media sosial mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Gillespie, 2018; Bucher, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi audit konten, platform diary, time-series engagement. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah strategi distribusi berintegritas, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.5. Rencana tindakan riset: Ekonomi Perhatian, Algoritma Rekomendasi, dan Dakwah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada media sosial; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi strategi distribusi berintegritas dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan audit konten, platform diary, time-series engagement.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur retensi sehat, meaningful engagement.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: clickbait agama, outrage amplification.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam ekonomi perhatian, algoritma rekomendasi, dan dakwah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah retensi sehat, meaningful engagement. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada media sosial, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus

dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Gillespie, 2018; Bucher, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan audit konten, platform diary, time-series engagement. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi strategi distribusi berintegritas: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada media sosial, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator retensi sehat, meaningful engagement dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa retensi sehat, meaningful engagement; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi clickbait agama, outrage amplification. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, strategi distribusi berintegritas memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks media sosial, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Gillespie, 2018; Bucher, 2018).

Contoh penerapan riil: pada media sosial, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas strategi distribusi berintegritas, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.5 Disinformasi, Misinformasi, dan Klaim Agama Viral

Disinformasi, Misinformasi, dan Klaim Agama Viral ditempatkan sebagai unit kerja riset karena membangun mekanisme deteksi dan koreksi klaim tanpa memperbesar penyebaran konten salah. Fokus pada komunitas digital mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Vosoughi et al., 2018; Wardle & Derakhshan, 2017. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi claim tracking, verification coding, source graph. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah protokol tabayyun digital, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.6. Rencana tindakan riset: Disinformasi, Misinformasi, dan Klaim Agama Viral

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada komunitas digital; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi protokol tabayyun digital dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan claim tracking, verification coding, source graph.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur precision verifikasi, correction uptake.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: backfire, amplifikasi rumor.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam disinformasi, misinformasi, dan klaim agama viral, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah precision verifikasi, correction uptake. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada komunitas digital, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus

dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Vosoughi et al., 2018; Wardle & Derakhshan, 2017 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan claim tracking, verification coding, source graph. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi protokol tabayyun digital: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada komunitas digital, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator precision verifikasi, correction uptake dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa precision verifikasi, correction uptake; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi backfire, amplifikasi rumor. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, protokol tabayyun digital memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks komunitas digital, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Vosoughi et al., 2018; Wardle & Derakhshan, 2017).

Contoh penerapan riil: pada komunitas digital, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas protokol tabayyun digital, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.6 Bahasa Lokal, Inklusivitas, dan Aksesibilitas

Bahasa Lokal, Inklusivitas, dan Aksesibilitas ditempatkan sebagai unit kerja riset karena mengembangkan layanan multi-bahasa dan akses bagi pengguna dengan variasi literasi serta disabilitas. Fokus pada masjid/sekolah mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021; Miao & Holmes, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi usability test, translation review, accessibility audit. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah materi multi-format, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.7. Rencana tindakan riset: Bahasa Lokal, Inklusivitas, dan Aksesibilitas

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid/sekolah; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi materi multi-format dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan usability test, translation review, accessibility audit.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur comprehension, WCAG-aligned accessibility.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: terjemahan salah nuansa, ketimpangan bahasa.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam bahasa lokal, inklusivitas, dan aksesibilitas, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah comprehension, WCAG-aligned accessibility. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid/sekolah, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus

dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021; Miao & Holmes, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan usability test, translation review, accessibility audit. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi materi multi-format: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid/sekolah, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator comprehension, WCAG-aligned accessibility dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa comprehension, WCAG-aligned accessibility; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi terjemahan salah nuansa, ketimpangan bahasa. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, materi multi-format memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid/sekolah, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021; Miao & Holmes, 2023).

Contoh penerapan riil: pada masjid/sekolah, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas materi multi-format, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.7 Digital Well-Being dan Batas Interaksi

Digital Well-Being dan Batas Interaksi ditempatkan sebagai unit kerja riset karena merancang dakwah digital yang tidak mendorong ketergantungan, overload notifikasi, atau interaksi manipulatif. Fokus pada remaja dan dewasa mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain OECD, 2024; UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi usage diary, well-being survey, focus group. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah kebijakan notifikasi dan desain sehat, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.8. Rencana tindakan riset: Digital Well-Being dan Batas Interaksi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei	Ringkasan	Masalah spesifik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	pada remaja dan dewasa; prioritaskan satu masalah.	baseline, consent, catatan sampling	dan terukur
2. Desain	Susun spesifikasi kebijakan notifikasi dan desain sehat dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan usage diary, well-being survey, focus group.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur retensi berkualitas, opt-out, well-being.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: dark patterns, endless engagement.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam digital well-being dan batas interaksi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah retensi berkualitas, opt-out, well-being. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada remaja dan dewasa, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. OECD, 2024; UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan usage diary, well-being survey, focus group. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi kebijakan notifikasi dan desain sehat: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada remaja dan dewasa, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim

mengumpulkan indikator retensi berkualitas, opt-out, well-being dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa retensi berkualitas, opt-out, well-being; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi dark patterns, endless engagement. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, kebijakan notifikasi dan desain sehat memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks remaja dan dewasa, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu

banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (OECD, 2024; UNESCO, 2021).

Contoh penerapan riil: pada remaja dan dewasa, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas kebijakan notifikasi dan desain sehat, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

2.8 Metode Riset Perilaku Digital yang Etis

Metode Riset Perilaku Digital yang Etis ditempatkan sebagai unit kerja riset karena memilih observasi, social listening, analytics, dan ethnography dengan perlindungan privasi dan informed consent. Fokus pada kampus dan lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UU 27/2022; Association of Internet Researchers, 2019. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi ethics protocol, data minimization. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah protokol penelitian digital, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 2.9. Rencana tindakan riset: Metode Riset Perilaku Digital yang Etis

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei	Ringkasan	Masalah spesifik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	pada kampus dan lembaga; prioritaskan satu masalah.	baseline, consent, catatan sampling	dan terukur
2. Desain	Susun spesifikasi protokol penelitian digital dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan ethics protocol, data minimization.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur consent coverage, de-identification quality.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: doxxing, re-identification, scraping berlebihan.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

2.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam metode riset perilaku digital yang etis, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah consent coverage, de-identification quality. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus dan lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi

sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UU 27/2022; Association of Internet Researchers, 2019 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

2.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan ethics protocol, data minimization. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi protokol penelitian digital: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus dan lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator consent coverage, de-identification quality dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem

'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

2.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa consent coverage, de-identification quality; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi doxxing, re-identification, scraping berlebihan. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, protokol penelitian digital memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus dan lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UU 27/2022; Association of Internet Researchers, 2019).

Contoh penerapan riil: pada kampus dan lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas protokol penelitian digital, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 2

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menganalisis hubungan praktik online dan offline serta perubahan komunitas, identitas, autentisitas, dan otoritas' sebagai contoh. Bukti dapat berasal dari digital ethnography, wawancara, diary study, sedangkan produk atau intervensinya dapat berbentuk peta perjalanan pengalaman beragama hibrida. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti keterlibatan lintas kanal, kualitas relasi harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti reduksi pengalaman ke klik; pengamatan invasif perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media sosial, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menelaah bagaimana visibilitas, reputasi, algoritma, dan jaringan memengaruhi siapa yang dianggap otoritatif' sebagai contoh. Bukti dapat berasal dari network analysis, content analysis, wawancara, sedangkan produk atau intervensinya dapat berbentuk peta otoritas dan mekanisme kredibilitas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti credential awareness, source checking harus dibaca bersama temuan kualitatif dan

catatan insiden. Risiko seperti pseudo-clerics, fatwa shopping, authority by virality perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, sekolah, kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengelompokkan kebutuhan audiens secara etis untuk penyesuaian materi tanpa mengeksploitasi kerentanan' sebagai contoh. Bukti dapat berasal dari survei, clustering anonim, persona research, sedangkan produk atau intervensinya dapat berbentuk persona kebutuhan dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti relevansi materi, akses, kepuasan harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling sensitif, discrimination perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media sosial, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami dampak ranking, engagement, dan insentif platform terhadap bentuk pesan keagamaan' sebagai contoh. Bukti dapat berasal dari audit konten, platform diary, time-series engagement, sedangkan produk atau intervensinya dapat berbentuk strategi distribusi berintegritas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retensi sehat, meaningful engagement harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti clickbait agama, outrage amplification perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas digital, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun mekanisme deteksi dan koreksi klaim tanpa memperbesar penyebaran konten salah' sebagai contoh. Bukti dapat berasal dari claim tracking, verification coding, source graph, sedangkan produk atau intervensinya dapat berbentuk protokol tabayyun digital. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti precision verifikasi,

correction uptake harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti backfire, amplifikasi rumor perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan layanan multi-bahasa dan akses bagi pengguna dengan variasi literasi serta disabilitas' sebagai contoh. Bukti dapat berasal dari usability test, translation review, accessibility audit, sedangkan produk atau intervensinya dapat berbentuk materi multi-format. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti comprehension, WCAG-aligned accessibility harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti terjemahan salah nuansa, ketimpangan bahasa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks remaja dan dewasa, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'merancang dakwah digital yang tidak mendorong ketergantungan, overload notifikasi, atau interaksi manipulatif' sebagai contoh. Bukti dapat berasal dari usage diary, well-being survey, focus group, sedangkan produk atau intervensinya dapat berbentuk kebijakan notifikasi dan desain sehat. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retensi berkualitas, opt-out, well-being harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dark patterns, endless engagement perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus dan lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memilih observasi, social listening, analytics, dan ethnography dengan perlindungan privasi dan informed consent' sebagai contoh. Bukti dapat berasal dari ethics protocol, data minimization, sedangkan produk atau intervensinya dapat berbentuk protokol penelitian digital. Kualitas tidak

cukup dinilai dari penggunaan atau popularitas; indikator seperti consent coverage, de-identification quality harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti doxxing, re-identification, scraping berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menganalisis hubungan praktik online dan offline serta perubahan komunitas, identitas, autentisitas, dan otoritas' sebagai contoh. Bukti dapat berasal dari digital ethnography, wawancara, diary study, sedangkan produk atau intervensinya dapat berbentuk peta perjalanan pengalaman beragama hibrida. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti keterlibatan lintas kanal, kualitas relasi harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti reduksi pengalaman ke klik; pengamatan invasif perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media sosial, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menelaah bagaimana visibilitas, reputasi, algoritma, dan jaringan memengaruhi siapa yang dianggap otoritatif' sebagai contoh. Bukti dapat berasal dari network analysis, content analysis, wawancara, sedangkan produk atau intervensinya dapat berbentuk peta otoritas dan mekanisme kredibilitas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti credential awareness, source checking harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pseudo-clerics, fatwa shopping, authority by virality perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, sekolah, kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengelompokkan kebutuhan audiens secara etis untuk penyesuaian materi tanpa mengeksploitasi kerentanan' sebagai contoh.

Bukti dapat berasal dari survei, clustering anonim, persona research, sedangkan produk atau intervensinya dapat berbentuk persona kebutuhan dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti relevansi materi, akses, kepuasan harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling sensitif, discrimination perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media sosial, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami dampak ranking, engagement, dan insentif platform terhadap bentuk pesan keagamaan' sebagai contoh. Bukti dapat berasal dari audit konten, platform diary, time-series engagement, sedangkan produk atau intervensinya dapat berbentuk strategi distribusi berintegritas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retensi sehat, meaningful engagement harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti clickbait agama, outrage amplification perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas digital, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun mekanisme deteksi dan koreksi klaim tanpa memperbesar penyebaran konten salah' sebagai contoh. Bukti dapat berasal dari claim tracking, verification coding, source graph, sedangkan produk atau intervensinya dapat berbentuk protokol tabayyun digital. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti precision verifikasi, correction uptake harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti backfire, amplifikasi rumor perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan layanan multi-bahasa dan akses bagi pengguna dengan variasi literasi serta disabilitas' sebagai contoh. Bukti dapat berasal dari usability test, translation review, accessibility audit, sedangkan

produk atau intervensinya dapat berbentuk materi multi-format. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti comprehension, WCAG-aligned accessibility harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti terjemahan salah nuansa, ketimpangan bahasa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks remaja dan dewasa, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'merancang dakwah digital yang tidak mendorong ketergantungan, overload notifikasi, atau interaksi manipulatif' sebagai contoh. Bukti dapat berasal dari usage diary, well-being survey, focus group, sedangkan produk atau intervensinya dapat berbentuk kebijakan notifikasi dan desain sehat. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retensi berkualitas, opt-out, well-being harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dark patterns, endless engagement perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus dan lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memilih observasi, social listening, analytics, dan ethnography dengan perlindungan privasi dan informed consent' sebagai contoh. Bukti dapat berasal dari ethics protocol, data minimization, sedangkan produk atau intervensinya dapat berbentuk protokol penelitian digital. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti consent coverage, de-identification quality harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti doxxing, re-identification, scraping berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas, tujuan yang terukur, data yang proporsional, dan kontrol

risiko. Gunakan fokus 'menganalisis hubungan praktik online dan offline serta perubahan komunitas, identitas, autentisitas, dan otoritas' sebagai contoh. Bukti dapat berasal dari digital ethnography, wawancara, diary study, sedangkan produk atau intervensinya dapat berbentuk peta perjalanan pengalaman beragama hibrida. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti keterlibatan lintas kanal, kualitas relasi harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti reduksi pengalaman ke klik; pengamatan invasif perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media sosial, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menelaah bagaimana visibilitas, reputasi, algoritma, dan jaringan memengaruhi siapa yang dianggap otoritatif' sebagai contoh. Bukti dapat berasal dari network analysis, content analysis, wawancara, sedangkan produk atau intervensinya dapat berbentuk peta otoritas dan mekanisme kredibilitas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti credential awareness, source checking harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pseudo-clerics, fatwa shopping, authority by virality perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, sekolah, kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengelompokkan kebutuhan audiens secara etis untuk penyesuaian materi tanpa mengeksploitasi kerentanan' sebagai contoh. Bukti dapat berasal dari survei, clustering anonim, persona research, sedangkan produk atau intervensinya dapat berbentuk persona kebutuhan dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti relevansi materi, akses, kepuasan harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling sensitif, discrimination perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media sosial, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memahami dampak ranking, engagement, dan insentif platform terhadap bentuk pesan keagamaan' sebagai contoh. Bukti dapat berasal dari audit konten, platform diary, time-series engagement, sedangkan produk atau intervensinya dapat berbentuk strategi distribusi berintegritas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retensi sehat, meaningful engagement harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti clickbait agama, outrage amplification perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

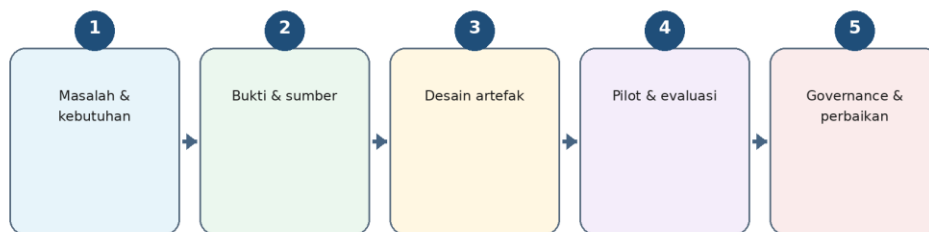
SOAL PENGEMBANGAN BAB 2

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 2 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 2, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 3 — METODOLOGI RISET DAN PENGEMBANGAN DAKWAH AI

Bab 3 membangun kompetensi untuk metodologi riset dan pengembangan dakwah ai. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 3



Gambar 3.1. Kerangka operasional Metodologi Riset Dan Pengembangan Dakwah Ai

Tabel 3.1. Matriks kerja utama Bab 3

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Desain	Bagaimana solusi bekerja dan di mana batasnya?	Spesifikasi + korpus	Prototipe	Uji internal
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

3.1 Merumuskan Masalah dan Theory of Change

Merumuskan Masalah dan Theory of Change ditempatkan sebagai unit kerja riset karena mengubah keluhan umum menjadi pertanyaan penelitian, mekanisme perubahan, outcome, dan asumsi yang dapat diuji. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Creswell & Creswell, 2018; Rogers, 2003. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi problem tree, stakeholder workshop, baseline. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah theory of change dakwah, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.2. Rencana tindakan riset: Merumuskan Masalah dan Theory of Change

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi theory of change dakwah dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan problem tree, stakeholder workshop, baseline.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur validitas konstruk, outcome clarity.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: solutionism dan target kabur.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam merumuskan masalah dan theory of change, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah validitas konstruk, outcome clarity. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Creswell & Creswell, 2018; Rogers, 2003 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan problem tree, stakeholder workshop, baseline. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi theory of change dakwah: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim

mengumpulkan indikator validitas konstruk, outcome clarity dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa validitas konstruk, outcome clarity; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi solutionism dan target kabur. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, theory of change dakwah memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna

menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Creswell & Creswell, 2018; Rogers, 2003).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas theory of change dakwah, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.2 Design Science Research untuk Artefak Dakwah

Design Science Research untuk Artefak Dakwah ditempatkan sebagai unit kerja riset karena mengembangkan dan mengevaluasi artefak digital melalui identifikasi masalah, tujuan, desain, demonstrasi, evaluasi, dan komunikasi. Fokus pada kampus/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Peffers et al., 2007; Hevner et al., 2004. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi DSRM. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah prototipe dan design principles, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.3. Rencana tindakan riset: Design Science Research untuk Artefak Dakwah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus/lembaga;	Ringkasan baseline,	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	prioritaskan satu masalah.	consent, catatan sampling	
2. Desain	Susun spesifikasi prototipe dan design principles dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan DSRM.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur utility, quality, efficacy.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: demo tanpa evaluasi; novelty bias.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam design science research untuk artefak dakwah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah utility, quality, efficacy. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan

meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Peffers et al., 2007; Hevner et al., 2004 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan DSRM. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi prototipe dan design principles: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator utility, quality, efficacy dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa utility, quality, efficacy; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi demo tanpa evaluasi; novelty bias. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, prototipe dan design principles memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Peffer et al., 2007; Hevner et al., 2004).

Contoh penerapan riil: pada kampus/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas prototipe dan design principles, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.3 Educational Design Research untuk Pembelajaran Agama

Educational Design Research untuk Pembelajaran Agama ditempatkan sebagai unit kerja riset karena mengembangkan intervensi pembelajaran melalui siklus analisis, desain, implementasi, dan refleksi di konteks nyata. Fokus pada sekolah/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain McKenney & Reeves, 2019; DBR Collective, 2003. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi EDR/DBR. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah learning design dan prinsip desain, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.4. Rencana tindakan riset: Educational Design Research untuk Pembelajaran Agama

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
2. Desain	Susun spesifikasi learning design dan prinsip desain dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan EDR/DBR.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur learning gain, transfer, usability.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: overfitting konteks, implementasi lemah.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam educational design research untuk pembelajaran agama, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah learning gain, transfer, usability. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. McKenney & Reeves, 2019; DBR Collective, 2003 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan EDR/DBR. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi learning design dan prinsip desain: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator learning gain, transfer, usability dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa learning gain, transfer, usability; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi overfitting konteks, implementasi lemah. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, learning design dan prinsip desain memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (McKenney & Reeves, 2019; DBR Collective, 2003).

Contoh penerapan riil: pada sekolah/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas learning design dan prinsip desain, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.4 Participatory Action Research untuk Komunitas

Participatory Action Research untuk Komunitas ditempatkan sebagai unit kerja riset karena melibatkan jamaah, dai, pengurus, guru, dan mahasiswa sebagai mitra peneliti dalam siklus perencanaan-tindakan-observasi-refleksi. Fokus pada masjid/komunitas mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Kemmis et al., 2014; Kemmis, 2009. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi PAR. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah rencana aksi komunitas, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.5. Rencana tindakan riset: Participatory Action Research untuk Komunitas

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid/komunitas; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
2. Desain	Susun spesifikasi rencana aksi komunitas dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan PAR.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur partisipasi, ownership, perubahan praktik.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: tokenisme partisipasi.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam participatory action research untuk komunitas, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah partisipasi, ownership, perubahan praktik. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid/komunitas, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Kemmis et al., 2014; Kemmis, 2009 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan PAR. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi rencana aksi komunitas: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid/komunitas, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator partisipasi, ownership, perubahan praktik dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa partisipasi, ownership, perubahan praktik; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi tokenisme partisipasi. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, rencana aksi komunitas memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid/komunitas, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Kemmis et al., 2014; Kemmis, 2009).

Contoh penerapan riil: pada masjid/komunitas, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas rencana aksi komunitas, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.5 Mixed Methods dan Triangulasi

Mixed Methods dan Triangulasi ditempatkan sebagai unit kerja riset karena menggabungkan data kuantitatif, kualitatif, log sistem, dan artefak untuk memperoleh penjelasan yang lebih utuh. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Creswell & Plano Clark, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi sequential/explanatory/convergent mixed methods. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah desain triangulasi, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.6. Rencana tindakan riset: Mixed Methods dan Triangulasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi desain	Spesifikasi,	Disetujui tim

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	triangulasi dan batas fungsi.	daftar sumber, risk register	substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan sequential/explanatory/convergent mixed methods.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur convergence, complementarity.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: data tidak terintegrasi.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam mixed methods dan triangulasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah convergence, complementarity. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah

pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Creswell & Plano Clark, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan sequential/explanatory/convergent mixed methods. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi desain triangulasi: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator convergence, complementarity dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa convergence, complementarity; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi data tidak terintegrasi. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, desain triangulasi memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Creswell & Plano Clark, 2018).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas desain triangulasi, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.6 Systematic Review, Evidence Mapping, dan PRISMA

Systematic Review, Evidence Mapping, dan PRISMA ditempatkan sebagai unit kerja riset karena menyusun tinjauan literatur transparan untuk menentukan state of the art, celah, serta komponen intervensi. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Page et al., 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi PRISMA 2020. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah evidence map, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.7. Rencana tindakan riset: Systematic Review, Evidence Mapping, dan PRISMA

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi evidence map dan batas fungsi.	Spesifikasi, daftar sumber,	Disetujui tim dan substansi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		risk register	teknis
3. Prototipe	Bangun versi minimum menggunakan PRISMA 2020.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur coverage database, screening reliability.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: selection bias, duplicate evidence.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam systematic review, evidence mapping, dan prisma, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah coverage database, screening reliability. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Page et al., 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan PRISMA 2020. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi evidence map: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator coverage database, screening reliability dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena

R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa coverage database, screening reliability; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi selection bias, duplicate evidence. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, evidence map memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Page et al., 2021).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-

keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas evidence map, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.7 Instrumen, Validitas, Reliabilitas, dan Uji Coba

Instrumen, Validitas, Reliabilitas, dan Uji Coba ditempatkan sebagai unit kerja riset karena mengembangkan kuesioner, rubrik, tes pengetahuan, interview guide, dan benchmark AI yang terukur. Fokus pada sekolah/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain DeVellis, 2017; Dillman et al., 2014. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi expert validation, pilot, reliability analysis. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah paket instrumen, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.8. Rencana tindakan riset: Instrumen, Validitas, Reliabilitas, dan Uji Coba

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi paket instrumen dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan expert validation, pilot, reliability	Version log, test set, defect log	Lulus uji internal minimum

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	analysis.		
4. Pilot	Uji terbatas; ukur content validity, reliability, inter-rater agreement.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: indikator ambigu, response bias.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam instrumen, validitas, reliabilitas, dan uji coba, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah content validity, reliability, inter-rater agreement. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. DeVellis, 2017; Dillman et al., 2014 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi,

kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan expert validation, pilot, reliability analysis. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi paket instrumen: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator content validity, reliability, inter-rater agreement dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa content validity, reliability, inter-rater agreement; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi indikator ambigu, response bias. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, paket instrumen memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (DeVellis, 2017; Dillman et al., 2014).

Contoh penerapan riil: pada sekolah/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas paket instrumen, dokumentasi

sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

3.8 Analisis Data dan Pelaporan yang Dapat Direplikasi

Analisis Data dan Pelaporan yang Dapat Direplikasi ditempatkan sebagai unit kerja riset karena merancang data dictionary, codebook, versioning, statistical plan, dan audit trail. Fokus pada kampus/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Creswell & Creswell, 2018; Braun & Clarke, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi analysis plan, reproducible workflow. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah research compendium, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 3.9. Rencana tindakan riset: Analisis Data dan Pelaporan yang Dapat Direplikasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus/lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi research compendium dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan analysis plan, reproducible workflow.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur	Dataset pilot,	Data cukup untuk

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	reproducibility, completeness.	log insiden, wawancara	keputusan
5. Keputusan	Tinjau risiko: p-hacking, selective reporting.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

3.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam analisis data dan pelaporan yang dapat direplikasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah reproducibility, completeness. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Creswell & Creswell, 2018; Braun & Clarke, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang

menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

3.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan analysis plan, reproducible workflow. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi research compendium: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator reproducibility, completeness dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

3.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa reproducibility, completeness; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi p-hacking, selective reporting. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, research compendium memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Creswell & Creswell, 2018; Braun & Clarke, 2021).

Contoh penerapan riil: pada kampus/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas research compendium, dokumentasi

sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 3

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengubah keluhan umum menjadi pertanyaan penelitian, mekanisme perubahan, outcome, dan asumsi yang dapat diuji' sebagai contoh. Bukti dapat berasal dari problem tree, stakeholder workshop, baseline, sedangkan produk atau intervensinya dapat berbentuk theory of change dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti validitas konstruk, outcome clarity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti solutionism dan target kabur perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan dan mengevaluasi artefak digital melalui identifikasi masalah, tujuan, desain, demonstrasi, evaluasi, dan komunikasi' sebagai contoh. Bukti dapat berasal dari DSRM, sedangkan produk atau intervensinya dapat berbentuk prototipe dan design principles. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti utility, quality, efficacy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti demo tanpa evaluasi; novelty bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan intervensi pembelajaran melalui siklus analisis, desain, implementasi, dan refleksi di konteks nyata' sebagai contoh. Bukti dapat berasal dari EDR/DBR, sedangkan produk atau intervensinya dapat

berbentuk learning design dan prinsip desain. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, transfer, usability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti overfitting konteks, implementasi lemah perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'melibatkan jamaah, dai, pengurus, guru, dan mahasiswa sebagai mitra peneliti dalam siklus perencanaan-tindakan-observasi-refleksi' sebagai contoh. Bukti dapat berasal dari PAR, sedangkan produk atau intervensinya dapat berbentuk rencana aksi komunitas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti partisipasi, ownership, perubahan praktik harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti tokenisme partisipasi perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggabungkan data kuantitatif, kualitatif, log sistem, dan artefak untuk memperoleh penjelasan yang lebih utuh' sebagai contoh. Bukti dapat berasal dari sequential/explanatory/convergent mixed methods, sedangkan produk atau intervensinya dapat berbentuk desain triangulasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti convergence, complementarity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data tidak terintegrasi perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun tinjauan literatur transparan untuk menentukan state of the art, celah, serta komponen intervensi' sebagai contoh. Bukti dapat berasal dari PRISMA 2020, sedangkan produk atau intervensinya dapat berbentuk evidence map. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator

seperti coverage database, screening reliability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti selection bias, duplicate evidence perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan kuesioner, rubrik, tes pengetahuan, interview guide, dan benchmark AI yang terukur' sebagai contoh. Bukti dapat berasal dari expert validation, pilot, reliability analysis, sedangkan produk atau intervensinya dapat berbentuk paket instrumen. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti content validity, reliability, inter-rater agreement harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti indikator ambigu, response bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'merancang data dictionary, codebook, versioning, statistical plan, dan audit trail' sebagai contoh. Bukti dapat berasal dari analysis plan, reproducible workflow, sedangkan produk atau intervensinya dapat berbentuk research compendium. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reproducibility, completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti p-hacking, selective reporting perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengubah keluhan umum menjadi pertanyaan penelitian, mekanisme perubahan, outcome, dan asumsi yang dapat diuji' sebagai contoh. Bukti dapat berasal dari problem tree, stakeholder workshop, baseline, sedangkan produk atau intervensinya dapat berbentuk theory of change dakwah. Kualitas

tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti validitas konstruk, outcome clarity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti solutionism dan target kabur perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan dan mengevaluasi artefak digital melalui identifikasi masalah, tujuan, desain, demonstrasi, evaluasi, dan komunikasi' sebagai contoh. Bukti dapat berasal dari DSRM, sedangkan produk atau intervensinya dapat berbentuk prototipe dan design principles. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti utility, quality, efficacy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti demo tanpa evaluasi; novelty bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan intervensi pembelajaran melalui siklus analisis, desain, implementasi, dan refleksi di konteks nyata' sebagai contoh. Bukti dapat berasal dari EDR/DBR, sedangkan produk atau intervensinya dapat berbentuk learning design dan prinsip desain. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, transfer, usability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti overfitting konteks, implementasi lemah perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'melibatkan jamaah, dai, pengurus, guru, dan mahasiswa sebagai mitra peneliti dalam siklus perencanaan-tindakan-observasi-refleksi' sebagai contoh. Bukti dapat berasal dari PAR, sedangkan produk atau intervensinya dapat berbentuk rencana aksi komunitas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti partisipasi, ownership, perubahan praktik harus dibaca bersama temuan kualitatif dan catatan insiden.

Risiko seperti tokenisme partisipasi perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggabungkan data kuantitatif, kualitatif, log sistem, dan artefak untuk memperoleh penjelasan yang lebih utuh' sebagai contoh. Bukti dapat berasal dari sequential/explanatory/convergent mixed methods, sedangkan produk atau intervensinya dapat berbentuk desain triangulasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti convergence, complementarity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data tidak terintegrasi perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun tinjauan literatur transparan untuk menentukan state of the art, celah, serta komponen intervensi' sebagai contoh. Bukti dapat berasal dari PRISMA 2020, sedangkan produk atau intervensinya dapat berbentuk evidence map. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti coverage database, screening reliability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti selection bias, duplicate evidence perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan kuesioner, rubrik, tes pengetahuan, interview guide, dan benchmark AI yang terukur' sebagai contoh. Bukti dapat berasal dari expert validation, pilot, reliability analysis, sedangkan produk atau intervensinya dapat berbentuk paket instrumen. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti content validity, reliability, inter-rater agreement harus dibaca bersama temuan kualitatif dan catatan insiden.

Risiko seperti indikator ambigu, response bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'merancang data dictionary, codebook, versioning, statistical plan, dan audit trail' sebagai contoh. Bukti dapat berasal dari analysis plan, reproducible workflow, sedangkan produk atau intervensinya dapat berbentuk research compendium. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reproducibility, completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti p-hacking, selective reporting perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengubah keluhan umum menjadi pertanyaan penelitian, mekanisme perubahan, outcome, dan asumsi yang dapat diuji' sebagai contoh. Bukti dapat berasal dari problem tree, stakeholder workshop, baseline, sedangkan produk atau intervensinya dapat berbentuk theory of change dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti validitas konstruk, outcome clarity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti solutionism dan target kabur perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan dan mengevaluasi artefak digital melalui identifikasi masalah, tujuan, desain, demonstrasi, evaluasi, dan komunikasi' sebagai contoh. Bukti dapat berasal dari DSRM, sedangkan produk atau intervensinya dapat berbentuk prototipe dan design principles. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti utility, quality, efficacy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti

demo tanpa evaluasi; novelty bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan intervensi pembelajaran melalui siklus analisis, desain, implementasi, dan refleksi di konteks nyata' sebagai contoh. Bukti dapat berasal dari EDR/DBR, sedangkan produk atau intervensinya dapat berbentuk learning design dan prinsip desain. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, transfer, usability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti overfitting konteks, implementasi lemah perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'melibatkan jamaah, dai, pengurus, guru, dan mahasiswa sebagai mitra peneliti dalam siklus perencanaan-tindakan-observasi-refleksi' sebagai contoh. Bukti dapat berasal dari PAR, sedangkan produk atau intervensinya dapat berbentuk rencana aksi komunitas. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti partisipasi, ownership, perubahan praktik harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti tokenisme partisipasi perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

SOAL PENGEMBANGAN BAB 3

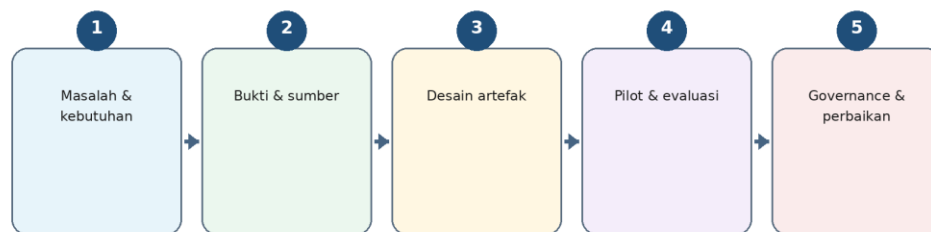
1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 3 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 3, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.

4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 4 — DATA, KORPUS, DAN KNOWLEDGE ENGINEERING DAKWAH

Bab 4 membangun kompetensi untuk data, korpus, dan knowledge engineering dakwah. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 4



Gambar 4.1. Kerangka operasional Data, Korpus, Dan Knowledge Engineering Dakwah

Tabel 4.1. Matriks kerja utama Bab 4

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Desain	Bagaimana solusi bekerja dan di mana batasnya?	Spesifikasi + korpus	Prototipe	Uji internal
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

4.1 Inventarisasi dan Kurasi Sumber Keagamaan

Inventarisasi dan Kurasi Sumber Keagamaan ditempatkan sebagai unit kerja riset karena membangun korpus dari sumber yang disetujui, dilengkapi metadata, edisi, otoritas, tema, dan status verifikasi. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Bunt, 2018; UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi document inventory, expert curation. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah korpus dakwah terverifikasi, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.2. Rencana tindakan riset: Inventarisasi dan Kurasi Sumber Keagamaan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi korpus dakwah terverifikasi dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan document inventory, expert curation.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur coverage, provenance completeness.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: sumber tidak sah, lisensi, versi campur.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam inventarisasi dan kurasi sumber keagamaan, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah coverage, provenance completeness. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Bunt, 2018; UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan document inventory, expert curation. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi korpus dakwah terverifikasi: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan

indikator coverage, provenance completeness dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa coverage, provenance completeness; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi sumber tidak sah, lisensi, versi campur. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, korpus dakwah terverifikasi memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna

menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Bunt, 2018; UNESCO, 2021).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas korpus dakwah terverifikasi, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.2 Taksonomi, Ontologi, dan Metadata Dakwah

Taksonomi, Ontologi, dan Metadata Dakwah ditempatkan sebagai unit kerja riset karena menyusun struktur tema aqidah, ibadah, akhlak, muamalah, keluarga, sosial, sejarah, dan kategori lembaga secara konsisten. Fokus pada lembaga/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Krippendorff, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi card sorting, ontology workshop. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah taksonomi/ontologi, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.3. Rencana tindakan riset: Taksonomi, Ontologi, dan Metadata Dakwah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		sampling	
2. Desain	Susun spesifikasi taksonomi/ontologi dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan card sorting, ontology workshop.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur inter-annotator agreement, retrieval precision.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: kategori bias dan tumpang tindih.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam taksonomi, ontologi, dan metadata dakwah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah inter-annotator agreement, retrieval precision. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Krippendorff, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan card sorting, ontology workshop. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi taksonomi/ontologi: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator inter-annotator agreement, retrieval precision dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa inter-annotator agreement, retrieval precision; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi kategori bias dan tumpang tindih. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, taksonomi/ontologi memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Krippendorff, 2018).

Contoh penerapan riil: pada lembaga/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas taksonomi/ontologi, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.3 Pembersihan, Normalisasi, dan Versioning Data

Pembersihan, Normalisasi, dan Versioning Data ditempatkan sebagai unit kerja riset karena menjaga teks, transliterasi, sumber, tanggal, bahasa, dan perubahan dokumen agar dapat ditelusuri. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2023; ISO/IEC 42001, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi ETL, deduplication, checksum. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah dataset versioned, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.4. Rencana tindakan riset: Pembersihan, Normalisasi, dan Versioning Data

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi dataset	Spesifikasi,	Disetujui tim

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	versioned dan batas fungsi.	daftar sumber, risk register	substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan ETL, deduplication, checksum.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur duplicate rate, error rate.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: silent modification, broken provenance.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam pembersihan, normalisasi, dan versioning data, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah duplicate rate, error rate. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2023; ISO/IEC 42001, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan ETL, deduplication, checksum. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi dataset versioned: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator duplicate rate, error rate dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena

R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa duplicate rate, error rate; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi silent modification, broken provenance. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, dataset versioned memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2023; ISO/IEC 42001, 2023).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara,

analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas dataset versioned, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.4 Embedding dan Pencarian Semantik

Embedding dan Pencarian Semantik ditempatkan sebagai unit kerja riset karena menggunakan representasi vektor untuk menemukan bagian relevan tanpa menyamakan kemiripan semantik dengan kebenaran normatif. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Lewis et al., 2020. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi retrieval benchmark. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah semantic search, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.5. Rencana tindakan riset: Embedding dan Pencarian Semantik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi semantic search dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan retrieval benchmark.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur recall@k,	Dataset pilot,	Data cukup untuk

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	precision@k, MRR.	log insiden, wawancara	keputusan
5. Keputusan	Tinjau risiko: retrieval salah tapi meyakinkan.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam embedding dan pencarian semantik, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah recall@k, precision@k, MRR. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Lewis et al., 2020 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan retrieval benchmark. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi semantic search: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator recall@k, precision@k, MRR dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa recall@k, precision@k, MRR; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi retrieval salah tapi meyakinkan. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, semantic search memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Lewis et al., 2020).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas semantic search, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.5 Retrieval-Augmented Generation untuk Tanya Jawab

Retrieval-Augmented Generation untuk Tanya Jawab ditempatkan sebagai unit kerja riset karena menggabungkan retrieval dan generation agar jawaban

didasarkan pada dokumen yang dapat ditunjukkan kepada pengguna. Fokus pada masjid/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Lewis et al., 2020; Ji et al., 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi RAG pipeline + human review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah asisten tanya-jawab non-fatwa, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.6. Rencana tindakan riset: Retrieval-Augmented Generation untuk Tanya Jawab

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid/lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi asisten tanya-jawab non-fatwa dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan RAG pipeline + human review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur citation accuracy, groundedness, abstention.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: hallucination, source laundering.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam retrieval-augmented generation untuk tanya jawab, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah citation accuracy, groundedness, abstention. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Lewis et al., 2020; Ji et al., 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan RAG pipeline + human review. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi asisten tanya-jawab non-fatwa: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator citation accuracy, groundedness, abstention dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa citation accuracy, groundedness, abstention; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi hallucination, source laundering. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang

menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, asisten tanya-jawab non-fatwa memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Lewis et al., 2020; Ji et al., 2023).

Contoh penerapan riil: pada masjid/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas asisten tanya-jawab non-fatwa, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.6 Dataset Uji untuk Pertanyaan Keagamaan

Dataset Uji untuk Pertanyaan Keagamaan ditempatkan sebagai unit kerja riset karena membangun gold set berisi pertanyaan representatif, jawaban rujukan, sumber, tingkat kesulitan, serta kategori eskalasi. Fokus pada kampus/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024. Dalam buku ini, teknologi diperlakukan

sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi expert-authored benchmark. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah test set, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.7. Rencana tindakan riset: Dataset Uji untuk Pertanyaan Keagamaan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus/lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi test set dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan expert-authored benchmark.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur coverage, agreement, adversarial robustness.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: benchmark leakage, narrow coverage.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam dataset uji untuk pertanyaan keagamaan, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah coverage, agreement, adversarial robustness. Pemisahan ini

penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan expert-authored benchmark. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi test set: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak

evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator coverage, agreement, adversarial robustness dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa coverage, agreement, adversarial robustness; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi benchmark leakage, narrow coverage. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus

berganti, atau biaya layanan meningkat. Karena itu, test set memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024).

Contoh penerapan riil: pada kampus/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas test set, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.7 Data Minimization dan Pelindungan Privasi

Data Minimization dan Pelindungan Privasi ditempatkan sebagai unit kerja riset karena membatasi pengumpulan data jamaah pada informasi yang benar-benar diperlukan dan memiliki dasar pemrosesan yang jelas. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UU 27/2022; UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi privacy impact assessment. Data tersebut tidak otomatis harus semuanya dikumpulkan.

Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah data governance plan, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.8. Rencana tindakan riset: Data Minimization dan Pelindungan Privasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi data governance plan dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan privacy impact assessment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur data minimization, retention compliance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: kebocoran data, consent invalid.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam data minimization dan pelindungan privasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah data minimization, retention compliance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen,

percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UU 27/2022; UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan privacy impact assessment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi data governance plan: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang

berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator data minimization, retention compliance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa data minimization, retention compliance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi kebocoran data, consent invalid. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, data governance plan memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik

sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UU 27/2022; UNESCO, 2021).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas data governance plan, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

4.8 Knowledge Maintenance dan Pembaruan Berkala

Knowledge Maintenance dan Pembaruan Berkala ditempatkan sebagai unit kerja riset karena menetapkan pemilik konten, jadwal review, pencabutan sumber, changelog, dan mekanisme rollback. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain ISO/IEC 42001, 2023; NIST, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi content lifecycle management. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah knowledge governance, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 4.9. Rencana tindakan riset: Knowledge Maintenance dan Pembaruan Berkala

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi knowledge governance dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan content lifecycle management.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur freshness, review SLA, rollback success.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: stale source, version drift.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

4.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam knowledge maintenance dan pembaruan berkala, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah freshness, review SLA, rollback success. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. ISO/IEC 42001, 2023; NIST, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

4.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan content lifecycle management. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi knowledge governance: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan

indikator freshness, review SLA, rollback success dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

4.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa freshness, review SLA, rollback success; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi stale source, version drift. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, knowledge governance memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna

menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (ISO/IEC 42001, 2023; NIST, 2023).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas knowledge governance, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 4

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun korpus dari sumber yang disetujui, dilengkapi metadata, edisi, otoritas, tema, dan status verifikasi' sebagai contoh. Bukti dapat berasal dari document inventory, expert curation, sedangkan produk atau intervensinya dapat berbentuk korpus dakwah terverifikasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti coverage, provenance completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti sumber tidak sah, lisensi, versi campur perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun struktur tema aqidah, ibadah, akhlak, muamalah, keluarga, sosial, sejarah, dan kategori lembaga secara konsisten' sebagai contoh. Bukti dapat berasal dari card sorting, ontology workshop, sedangkan produk atau intervensinya dapat berbentuk taksonomi/ontologi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti inter-annotator agreement, retrieval precision harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kategori bias dan tumpang tindih perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke

kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjaga teks, transliterasi, sumber, tanggal, bahasa, dan perubahan dokumen agar dapat ditelusuri' sebagai contoh. Bukti dapat berasal dari ETL, deduplication, checksum, sedangkan produk atau intervensinya dapat berbentuk dataset versioned. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti duplicate rate, error rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti silent modification, broken provenance perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan representasi vektor untuk menemukan bagian relevan tanpa menyamakan kemiripan semantik dengan kebenaran normatif' sebagai contoh. Bukti dapat berasal dari retrieval benchmark, sedangkan produk atau intervensinya dapat berbentuk semantic search. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti recall@k, precision@k, MRR harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti retrieval salah tapi meyakinkan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggabungkan retrieval dan generation agar jawaban didasarkan pada dokumen yang dapat ditunjukkan kepada pengguna' sebagai contoh. Bukti dapat berasal dari RAG pipeline + human review, sedangkan produk atau intervensinya dapat berbentuk asisten tanya-jawab non-fatwa. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti citation accuracy, groundedness, abstention harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti hallucination, source laundering perlu

ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun gold set berisi pertanyaan representatif, jawaban rujukan, sumber, tingkat kesulitan, serta kategori eskalasi' sebagai contoh. Bukti dapat berasal dari expert-authored benchmark, sedangkan produk atau intervensinya dapat berbentuk test set. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti coverage, agreement, adversarial robustness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti benchmark leakage, narrow coverage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membatasi pengumpulan data jamaah pada informasi yang benar-benar diperlukan dan memiliki dasar pemrosesan yang jelas' sebagai contoh. Bukti dapat berasal dari privacy impact assessment, sedangkan produk atau intervensinya dapat berbentuk data governance plan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti data minimization, retention compliance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kebocoran data, consent invalid perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menetapkan pemilik konten, jadwal review, pencabutan sumber, changelog, dan mekanisme rollback' sebagai contoh. Bukti dapat berasal dari content lifecycle management, sedangkan produk atau intervensinya dapat berbentuk knowledge governance. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti freshness, review SLA, rollback success harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti stale source, version drift perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan

ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun korpus dari sumber yang disetujui, dilengkapi metadata, edisi, otoritas, tema, dan status verifikasi' sebagai contoh. Bukti dapat berasal dari document inventory, expert curation, sedangkan produk atau intervensinya dapat berbentuk korpus dakwah terverifikasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti coverage, provenance completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti sumber tidak sah, lisensi, versi campur perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun struktur tema aqidah, ibadah, akhlak, muamalah, keluarga, sosial, sejarah, dan kategori lembaga secara konsisten' sebagai contoh. Bukti dapat berasal dari card sorting, ontology workshop, sedangkan produk atau intervensinya dapat berbentuk taksonomi/ontologi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti inter-annotator agreement, retrieval precision harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kategori bias dan tumpang tindih perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjaga teks, transliterasi, sumber, tanggal, bahasa, dan perubahan dokumen agar dapat ditelusuri' sebagai contoh. Bukti dapat berasal dari ETL, deduplication, checksum, sedangkan produk atau intervensinya dapat berbentuk dataset versioned. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti duplicate rate, error rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti silent modification, broken provenance perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan

ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan representasi vektor untuk menemukan bagian relevan tanpa menyamakan kemiripan semantik dengan kebenaran normatif' sebagai contoh. Bukti dapat berasal dari retrieval benchmark, sedangkan produk atau intervensinya dapat berbentuk semantic search. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti recall@k, precision@k, MRR harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti retrieval salah tapi meyakinkan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggabungkan retrieval dan generation agar jawaban didasarkan pada dokumen yang dapat ditunjukkan kepada pengguna' sebagai contoh. Bukti dapat berasal dari RAG pipeline + human review, sedangkan produk atau intervensinya dapat berbentuk asisten tanya-jawab non-fatwa. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti citation accuracy, groundedness, abstention harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti hallucination, source laundering perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun gold set berisi pertanyaan representatif, jawaban rujukan, sumber, tingkat kesulitan, serta kategori eskalasi' sebagai contoh. Bukti dapat berasal dari expert-authored benchmark, sedangkan produk atau intervensinya dapat berbentuk test set. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti coverage, agreement, adversarial robustness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti benchmark leakage, narrow coverage perlu ditulis sebagai skenario

konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membatasi pengumpulan data jamaah pada informasi yang benar-benar diperlukan dan memiliki dasar pemrosesan yang jelas' sebagai contoh. Bukti dapat berasal dari privacy impact assessment, sedangkan produk atau intervensinya dapat berbentuk data governance plan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti data minimization, retention compliance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kebocoran data, consent invalid perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menetapkan pemilik konten, jadwal review, pencabutan sumber, changelog, dan mekanisme rollback' sebagai contoh. Bukti dapat berasal dari content lifecycle management, sedangkan produk atau intervensinya dapat berbentuk knowledge governance. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti freshness, review SLA, rollback success harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti stale source, version drift perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun korpus dari sumber yang disetujui, dilengkapi metadata, edisi, otoritas, tema, dan status verifikasi' sebagai contoh. Bukti dapat berasal dari document inventory, expert curation, sedangkan produk atau intervensinya dapat berbentuk korpus dakwah terverifikasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti coverage, provenance completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti sumber tidak sah, lisensi, versi campur perlu ditulis sebagai skenario

konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun struktur tema aqidah, ibadah, akhlak, muamalah, keluarga, sosial, sejarah, dan kategori lembaga secara konsisten' sebagai contoh. Bukti dapat berasal dari card sorting, ontology workshop, sedangkan produk atau intervensinya dapat berbentuk taksonomi/ontologi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti inter-annotator agreement, retrieval precision harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti kategori bias dan tumpang tindih perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjaga teks, transliterasi, sumber, tanggal, bahasa, dan perubahan dokumen agar dapat ditelusuri' sebagai contoh. Bukti dapat berasal dari ETL, deduplication, checksum, sedangkan produk atau intervensinya dapat berbentuk dataset versioned. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti duplicate rate, error rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti silent modification, broken provenance perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan representasi vektor untuk menemukan bagian relevan tanpa menyamakan kemiripan semantik dengan kebenaran normatif' sebagai contoh. Bukti dapat berasal dari retrieval benchmark, sedangkan produk atau intervensinya dapat berbentuk semantic search. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti recall@k, precision@k, MRR harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti

retrieval salah tapi meyakinkan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

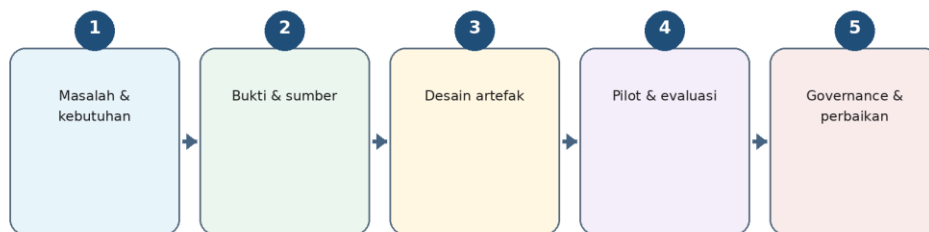
SOAL PENGEMBANGAN BAB 4

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 4 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 4, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 5 — PENGEMBANGAN MEDIA, CHATBOT, DAN KONTEN GENERATIF

Bab 5 membangun kompetensi untuk pengembangan media, chatbot, dan konten generatif. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 5



Gambar 5.1. Kerangka operasional Pengembangan Media, Chatbot, Dan Konten Generatif

Tabel 5.1. Matriks kerja utama Bab 5

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Desain	Bagaimana solusi bekerja dan di mana batasnya?	Spesifikasi + korpus	Prototipe	Uji internal
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

5.1 Prompt Engineering sebagai Spesifikasi Tugas

Prompt Engineering sebagai Spesifikasi Tugas ditempatkan sebagai unit kerja riset karena menulis prompt sebagai instruksi yang memuat tujuan, konteks, sumber, format, larangan, kriteria kualitas, dan kondisi abstain. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024; Miao & Holmes, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi prompt experiment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah prompt library terverifikasi, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.2. Rencana tindakan riset: Prompt Engineering sebagai Spesifikasi Tugas

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi prompt library terverifikasi dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan prompt experiment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur consistency, error rate.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: prompt injection, hidden assumptions.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam prompt engineering sebagai spesifikasi tugas, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah consistency, error rate. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024; Miao & Holmes, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan prompt experiment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi prompt library terverifikasi: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim

mengumpulkan indikator consistency, error rate dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa consistency, error rate; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi prompt injection, hidden assumptions. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, prompt library terverifikasi memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024; Miao & Holmes, 2023).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas prompt library terverifikasi, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.2 Chatbot Dakwah Non-Fatwa Berbasis RAG

Chatbot Dakwah Non-Fatwa Berbasis RAG ditempatkan sebagai unit kerja riset karena membangun chatbot untuk pencarian informasi, jadwal, materi dasar, dan rujukan dengan sumber eksplisit serta eskalasi manusia. Fokus pada masjid/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Lewis et al., 2020; Bunt, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi DSRM + RAG evaluation. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah chatbot layanan, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.3. Rencana tindakan riset: Chatbot Dakwah Non-Fatwa Berbasis RAG

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid/lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		sampling	
2. Desain	Susun spesifikasi chatbot layanan dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan DSRM + RAG evaluation.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur groundedness, citation precision, escalation.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: AI overreach, hallucination.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam chatbot dakwah non-fatwa berbasis rag, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah groundedness, citation precision, escalation. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Lewis et al., 2020; Bunt, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan DSRM + RAG evaluation. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi chatbot layanan: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator groundedness, citation precision, escalation dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa groundedness, citation precision, escalation; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi AI overreach, hallucination. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, chatbot layanan memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Lewis et al., 2020; Bunt, 2018).

Contoh penerapan riil: pada masjid/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas chatbot layanan, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.3 Generasi Ringkasan, FAQ, dan Microlearning

Generasi Ringkasan, FAQ, dan Microlearning ditempatkan sebagai unit kerja riset karena menggunakan AI untuk menghasilkan draft dari sumber yang telah ditetapkan lalu melewati review editorial. Fokus pada sekolah/masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Miao & Holmes, 2023; Kasneci et al., 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi controlled generation + editorial review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah paket microlearning, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.4. Rencana tindakan riset: Generasi Ringkasan, FAQ, dan Microlearning

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah/masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi paket	Spesifikasi,	Disetujui tim

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	microlearning dan batas fungsi.	daftar sumber, risk register	substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan controlled generation + editorial review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur factual fidelity, readability.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: distorsi makna, oversimplification.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam generasi ringkasan, faq, dan microlearning, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah factual fidelity, readability. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah/masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Miao & Holmes, 2023; Kasneci et al., 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan controlled generation + editorial review. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi paket microlearning: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah/masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator factual fidelity, readability dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila

kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa factual fidelity, readability; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi distorsi makna, oversimplification. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, paket microlearning memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah/masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Miao & Holmes, 2023; Kasneci et al., 2023).

Contoh penerapan riil: pada sekolah/masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang

dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas paket microlearning, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.4 Transkripsi, Caption, dan Terjemahan Ceramah

Transkripsi, Caption, dan Terjemahan Ceramah ditempatkan sebagai unit kerja riset karena mengubah audio/video menjadi teks, caption, indeks topik, dan terjemahan untuk memperluas akses. Fokus pada masjid/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi ASR evaluation + bilingual review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah arsip ceramah searchable, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.5. Rencana tindakan riset: Transkripsi, Caption, dan Terjemahan Ceramah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi arsip ceramah searchable dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
3. Prototipe	Bangun versi minimum menggunakan ASR evaluation + bilingual review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur WER, terminology accuracy.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: nama/istilah agama salah, privacy.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam transkripsi, caption, dan terjemahan ceramah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah WER, terminology accuracy. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan ASR evaluation + bilingual review. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi arsip ceramah searchable: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator WER, terminology accuracy dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena

R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa WER, terminology accuracy; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi nama/istilah agama salah, privacy. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, arsip ceramah searchable memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021).

Contoh penerapan riil: pada masjid/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot,

wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas arsip ceramah searchable, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.5 Visualisasi dan Desain Konten Berbantuan AI

Visualisasi dan Desain Konten Berbantuan AI ditempatkan sebagai unit kerja riset karena menggunakan AI untuk ide visual dan layout dengan prinsip tidak memalsukan tokoh, peristiwa, atau sumber. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024; EU AI Act, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi design review, provenance labeling. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah template visual dakwah, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.6. Rencana tindakan riset: Visualisasi dan Desain Konten Berbantuan AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi template visual dakwah dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan design review, provenance	Version log, test set, defect log	Lulus uji internal minimum

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	labeling.		
4. Pilot	Uji terbatas; ukur clarity, accessibility, provenance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: deepfake, synthetic misrepresentation.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam visualisasi dan desain konten berbantuan ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah clarity, accessibility, provenance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024; EU AI Act, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal

sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan design review, provenance labeling. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi template visual dakwah: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator clarity, accessibility, provenance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa clarity, accessibility, provenance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi deepfake, synthetic misrepresentation. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, template visual dakwah memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024; EU AI Act, 2024).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas template visual dakwah, dokumentasi sumber, instrumen

evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.6 Personalisasi yang Proporsional dan Transparan

Personalisasi yang Proporsional dan Transparan ditempatkan sebagai unit kerja riset karena menyesuaikan urutan materi berdasarkan pilihan eksplisit atau progres belajar tanpa membangun profil sensitif tersembunyi. Fokus pada sekolah/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain OECD, 2024; UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi adaptive learning pilot. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah learning pathway, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.7. Rencana tindakan riset: Personalisasi yang Proporsional dan Transparan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi learning pathway dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan adaptive learning pilot.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur learning gain, opt-out rate.	Dataset pilot, log insiden,	Data cukup untuk keputusan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		wawancara	
5. Keputusan	Tinjau risiko: profiling, filter bubble.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam personalisasi yang proporsional dan transparan, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah learning gain, opt-out rate. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. OECD, 2024; UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan adaptive learning pilot. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi learning pathway: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator learning gain, opt-out rate dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa learning gain, opt-out rate; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi profiling, filter bubble. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, learning pathway memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (OECD, 2024; UNESCO, 2021).

Contoh penerapan riil: pada sekolah/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas learning pathway, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.7 Quality Assurance Konten AI

Quality Assurance Konten AI ditempatkan sebagai unit kerja riset karena membuat pipeline pemeriksaan sumber, bahasa, konteks, kesesuaian audiens, etika, dan persetujuan penerbitan. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain ISO/IEC 42001, 2023; NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi multi-stage review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah QA checklist, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.8. Rencana tindakan riset: Quality Assurance Konten AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi QA checklist dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan multi-stage review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur defect rate, time-to-review.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: review fatigue, rubber stamping.	Minutes review, keputusan dan	Go/revise/hold/stop terdokumentasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		alasan	

5.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam quality assurance konten ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah defect rate, time-to-review. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. ISO/IEC 42001, 2023; NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan multi-stage review. Hasil diagnosis ditulis sebagai daftar kebutuhan

terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi QA checklist: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator defect rate, time-to-review dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa defect rate, time-to-review; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi review fatigue, rubber stamping. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak,

likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, QA checklist memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (ISO/IEC 42001, 2023; NIST, 2024).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas QA checklist, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

5.8 Publikasi, Provenance, dan Pelabelan Konten

Publikasi, Provenance, dan Pelabelan Konten ditempatkan sebagai unit kerja riset karena menjelaskan kapan AI digunakan, siapa editor manusia, sumber utama, tanggal review, dan versi materi. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021; OECD, 2024. Dalam buku ini, teknologi diperlakukan sebagai

komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi provenance protocol. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah label transparansi, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 5.9. Rencana tindakan riset: Publikasi, Provenance, dan Pelabelan Konten

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi label transparansi dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan provenance protocol.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur traceability, complaint resolution.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: misleading synthetic content.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

5.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam publikasi, provenance, dan pelabelan konten, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah traceability, complaint resolution. Pemisahan ini penting

agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021; OECD, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

5.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan provenance protocol. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi label transparansi: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator traceability, complaint resolution dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

5.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa traceability, complaint resolution; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi misleading synthetic content. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, label transparansi memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021; OECD, 2024).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas label transparansi, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 5

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menulis prompt sebagai instruksi yang memuat tujuan, konteks, sumber, format, larangan, kriteria kualitas, dan kondisi abstain' sebagai contoh. Bukti dapat berasal dari prompt experiment, sedangkan produk atau intervensinya dapat berbentuk prompt library terverifikasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti consistency, error rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti prompt injection, hidden assumptions perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun chatbot untuk pencarian informasi, jadwal, materi dasar, dan rujukan dengan sumber eksplisit serta eskalasi manusia' sebagai contoh. Bukti dapat berasal dari DSRM + RAG evaluation, sedangkan produk atau intervensinya dapat berbentuk chatbot layanan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti groundedness, citation precision, escalation harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI overreach, hallucination perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengggunakan AI untuk menghasilkan draft dari sumber yang telah ditetapkan lalu melewati review editorial' sebagai contoh. Bukti dapat berasal dari controlled generation + editorial review, sedangkan produk atau intervensinya dapat berbentuk paket microlearning. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti factual fidelity, readability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti distorsi makna, oversimplification perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengubah audio/video menjadi teks, caption, indeks topik, dan terjemahan untuk memperluas akses' sebagai contoh. Bukti dapat berasal dari ASR evaluation + bilingual review, sedangkan produk atau intervensinya dapat berbentuk arsip ceramah searchable. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti WER, terminology accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti nama/istilah agama salah, privacy perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan AI untuk ide visual dan layout dengan prinsip tidak memalsukan tokoh, peristiwa, atau sumber' sebagai contoh. Bukti dapat berasal dari design review, provenance labeling, sedangkan produk atau intervensinya dapat berbentuk template visual dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti clarity, accessibility, provenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti deepfake, synthetic misrepresentation perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyesuaikan urutan materi berdasarkan pilihan eksplisit atau progres belajar tanpa membangun profil sensitif tersembunyi' sebagai contoh. Bukti dapat berasal dari adaptive learning pilot, sedangkan produk atau intervensinya dapat berbentuk learning pathway. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, opt-out rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling, filter bubble perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat pipeline pemeriksaan sumber, bahasa, konteks, kesesuaian audiens, etika, dan persetujuan penerbitan' sebagai contoh. Bukti dapat berasal dari multi-stage review, sedangkan produk atau intervensinya dapat berbentuk QA checklist. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti defect rate, time-to-review harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti review fatigue, rubber stamping perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjelaskan kapan AI digunakan, siapa editor manusia, sumber utama, tanggal review, dan versi materi' sebagai contoh. Bukti dapat berasal dari provenance protocol, sedangkan produk atau intervensinya dapat berbentuk label transparansi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti traceability, complaint resolution harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti misleading synthetic content perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menulis prompt sebagai instruksi yang memuat tujuan, konteks, sumber, format, larangan, kriteria kualitas, dan kondisi abstain' sebagai contoh. Bukti dapat berasal dari prompt experiment, sedangkan produk atau intervensinya dapat berbentuk prompt library terverifikasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti consistency, error rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti prompt injection, hidden assumptions perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun chatbot untuk pencarian informasi, jadwal, materi dasar, dan rujukan dengan sumber eksplisit serta eskalasi manusia' sebagai contoh. Bukti dapat berasal dari DSRM + RAG evaluation, sedangkan produk atau intervensinya dapat berbentuk chatbot layanan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti groundedness, citation precision, escalation harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI overreach, hallucination perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan AI untuk menghasilkan draft dari sumber yang telah ditetapkan lalu melewati review editorial' sebagai contoh. Bukti dapat berasal dari controlled generation + editorial review, sedangkan produk atau intervensinya dapat berbentuk paket microlearning. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti factual fidelity, readability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti distorsi makna, oversimplification perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengubah audio/video menjadi teks, caption, indeks topik, dan terjemahan untuk memperluas akses' sebagai contoh. Bukti dapat berasal dari ASR evaluation + bilingual review, sedangkan produk atau intervensinya dapat berbentuk arsip ceramah searchable. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti WER, terminology accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti nama/istilah agama salah, privacy perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan AI untuk ide visual dan layout dengan prinsip tidak memalsukan tokoh, peristiwa, atau sumber' sebagai contoh. Bukti dapat berasal dari design review, provenance labeling, sedangkan produk atau intervensinya dapat berbentuk template visual dakwah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti clarity, accessibility, provenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti deepfake, synthetic misrepresentation perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyesuaikan urutan materi berdasarkan pilihan eksplisit atau progres belajar tanpa membangun profil sensitif tersembunyi' sebagai contoh. Bukti dapat berasal dari adaptive learning pilot, sedangkan produk atau intervensinya dapat berbentuk learning pathway. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, opt-out rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling, filter bubble perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat pipeline pemeriksaan sumber, bahasa, konteks, kesesuaian audiens, etika, dan persetujuan penerbitan' sebagai contoh. Bukti dapat berasal dari multi-stage review, sedangkan produk atau intervensinya dapat berbentuk QA checklist. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti defect rate, time-to-review harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti review fatigue, rubber stamping perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjelaskan kapan AI digunakan, siapa editor manusia, sumber utama, tanggal review, dan versi materi' sebagai contoh. Bukti dapat berasal dari provenance protocol, sedangkan produk atau intervensinya dapat berbentuk label transparansi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti traceability, complaint resolution harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti misleading synthetic content perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menulis prompt sebagai instruksi yang memuat tujuan, konteks, sumber, format, larangan, kriteria kualitas, dan kondisi abstain' sebagai contoh. Bukti dapat berasal dari prompt experiment, sedangkan produk atau intervensinya dapat berbentuk prompt library terverifikasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti consistency, error rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti prompt injection, hidden assumptions perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun chatbot untuk pencarian informasi, jadwal, materi dasar, dan rujukan dengan sumber eksplisit serta eskalasi manusia' sebagai contoh. Bukti dapat berasal dari DSRM + RAG evaluation, sedangkan produk atau intervensinya dapat berbentuk chatbot layanan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti groundedness, citation precision, escalation harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI overreach, hallucination perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan AI untuk menghasilkan draft dari sumber yang telah ditetapkan lalu melewati review editorial' sebagai contoh. Bukti dapat berasal dari controlled generation + editorial review, sedangkan produk atau intervensinya dapat berbentuk paket microlearning. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti factual fidelity, readability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti distorsi makna, oversimplification perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengubah audio/video menjadi teks, caption, indeks topik, dan terjemahan untuk memperluas akses' sebagai contoh. Bukti dapat berasal dari ASR evaluation + bilingual review, sedangkan produk atau intervensinya dapat berbentuk arsip ceramah searchable. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti WER, terminology accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti nama/istilah agama salah, privacy perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

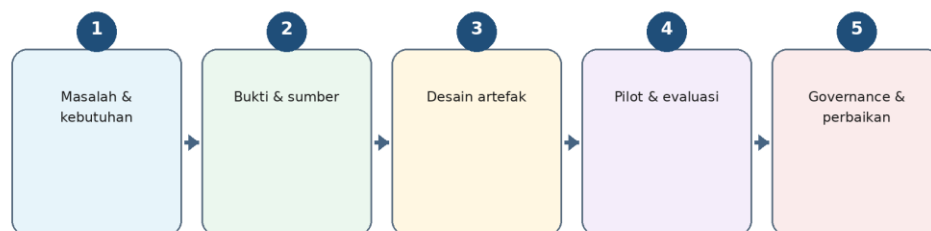
SOAL PENGEMBANGAN BAB 5

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 5 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 5, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 6 — IMPLEMENTASI RIIL DI MASJID DAN KOMUNITAS

Bab 6 membangun kompetensi untuk implementasi riil di masjid dan komunitas. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 6



Gambar 6.1. Kerangka operasional Implementasi Riil Di Masjid Dan Komunitas

Tabel 6.1. Matriks kerja utama Bab 6

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan
Desain	Bagaimana solusi bekerja	Spesifikasi + korpus	Prototipe	Uji internal

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
	dan di mana batasnya?			
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

6.1 Audit Kebutuhan Jamaah 30 Hari

Audit Kebutuhan Jamaah 30 Hari ditempatkan sebagai unit kerja riset karena menjalankan diagnosis cepat masalah layanan, pembelajaran, komunikasi, akses, dan kebutuhan kelompok jamaah. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Creswell & Creswell, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi survei singkat, wawancara, observasi. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah baseline kebutuhan jamaah, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.2. Rencana tindakan riset: Audit Kebutuhan Jamaah 30 Hari

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi baseline kebutuhan jamaah dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan survei singkat, wawancara, observasi.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur response coverage, priority score.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: survei bias, data berlebih.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam audit kebutuhan jamaah 30 hari, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah response coverage, priority score. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Creswell & Creswell, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan survei singkat, wawancara, observasi. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi baseline kebutuhan jamaah: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan

indikator response coverage, priority score dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa response coverage, priority score; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi survei bias, data berlebih. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, baseline kebutuhan jamaah memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Creswell & Creswell, 2018).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas baseline kebutuhan jamaah, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.2 Pusat Pengetahuan Masjid Terverifikasi

Pusat Pengetahuan Masjid Terverifikasi ditempatkan sebagai unit kerja riset karena mengonsolidasikan SOP, jadwal, materi kajian, program sosial, panduan ibadah institusional, dan kontak layanan dalam basis pengetahuan. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Lewis et al., 2020; ISO/IEC 42001, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi content inventory + RAG. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah knowledge hub, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.3. Rencana tindakan riset: Pusat Pengetahuan Masjid Terverifikasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
2. Desain	Susun spesifikasi knowledge hub dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan content inventory + RAG.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur search success, freshness.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: dokumen kedaluwarsa.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam pusat pengetahuan masjid terverifikasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah search success, freshness. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Lewis et al., 2020; ISO/IEC 42001, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan content inventory + RAG. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi knowledge hub: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator search success, freshness dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila

kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa search success, freshness; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi dokumen kedaluwarsa. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, knowledge hub memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Lewis et al., 2020; ISO/IEC 42001, 2023).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline

dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas knowledge hub, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.3 Asisten Layanan Jamaah Non-Fatwa

Asisten Layanan Jamaah Non-Fatwa ditempatkan sebagai unit kerja riset karena memberi jawaban tentang jadwal, lokasi, pendaftaran kegiatan, materi rujukan, dan FAQ administrasi dengan eskalasi manusia. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi service design + chatbot pilot. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah service assistant, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.4. Rencana tindakan riset: Asisten Layanan Jamaah Non-Fatwa

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi service assistant dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan service design + chatbot pilot.	Version log, test set, defect log	Lulus uji internal minimum

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
4. Pilot	Uji terbatas; ukur resolution rate, escalation accuracy.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: pertanyaan hukum agama masuk scope.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam asisten layanan jamaah non-fatwa, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah resolution rate, escalation accuracy. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan,

mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan service design + chatbot pilot. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi service assistant: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator resolution rate, escalation accuracy dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa resolution rate, escalation accuracy; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi pertanyaan hukum agama masuk scope. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, service assistant memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas service assistant, dokumentasi sumber, instrumen

evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.4 Arsip Khutbah dan Kajian yang Dapat Dicari

Arsip Khutbah dan Kajian yang Dapat Dicari ditempatkan sebagai unit kerja riset karena mentranskripsi, memberi metadata, ringkasan, tema, dan tautan sumber untuk membangun memori institusional. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Campbell & Bellar, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi ASR + curation. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah digital sermon archive, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.5. Rencana tindakan riset: Arsip Khutbah dan Kajian yang Dapat Dicari

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi digital sermon archive dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan ASR + curation.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur searchability, transcript	Dataset pilot, log insiden,	Data cukup untuk keputusan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	accuracy.	wawancara	
5. Keputusan	Tinjau risiko: copyright, misquote.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam arsip khutbah dan kajian yang dapat dicari, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah searchability, transcript accuracy. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Campbell & Bellar, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan ASR + curation. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi digital sermon archive: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator searchability, transcript accuracy dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa searchability, transcript accuracy; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi copyright, misquote. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, digital sermon archive memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Campbell & Bellar, 2023).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas digital sermon archive, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.5 Dakwah Berbasis Data Kebutuhan Sosial

Dakwah Berbasis Data Kebutuhan Sosial ditempatkan sebagai unit kerja riset karena menghubungkan program dakwah dengan data kebutuhan pendidikan,

kemiskinan, kesehatan sosial, keluarga, dan relawan secara agregat. Fokus pada komunitas mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UU 27/2022; RE-AIM. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi needs assessment + program evaluation. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah dashboard program sosial, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.6. Rencana tindakan riset: Dakwah Berbasis Data Kebutuhan Sosial

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada komunitas; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi dashboard program sosial dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan needs assessment + program evaluation.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur reach, referral, completion.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: stigma, sensitive data.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam dakwah berbasis data kebutuhan sosial, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah reach, referral, completion. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada komunitas, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UU 27/2022; RE-AIM memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan needs assessment + program evaluation. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi dashboard program sosial: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada komunitas, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator reach, referral, completion dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa reach, referral, completion; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi stigma, sensitive data. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, dashboard program sosial memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks komunitas, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UU 27/2022; RE-AIM).

Contoh penerapan riil: pada komunitas, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas dashboard program sosial, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.6 Program Literasi AI Jamaah dan Keluarga

Program Literasi AI Jamaah dan Keluarga ditempatkan sebagai unit kerja riset karena mendidik jamaah untuk memeriksa sumber, mengenali konten sintetis, menjaga privasi, dan menggunakan AI secara proporsional. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Miao & Holmes, 2023; NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi pre-post workshop. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah modul literasi, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.7. Rencana tindakan riset: Program Literasi AI Jamaah dan Keluarga

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi modul literasi dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan pre-post workshop.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur knowledge gain, detection accuracy.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: fear-mongering atau overconfidence.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam program literasi ai jamaah dan keluarga, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah knowledge gain, detection accuracy. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki

definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Miao & Holmes, 2023; NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan pre-post workshop. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi modul literasi: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator knowledge gain, detection accuracy dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa knowledge gain, detection accuracy; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi fear-mongering atau overconfidence. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, modul literasi memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Miao & Holmes, 2023; NIST, 2024).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas modul literasi, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.7 Volunteer Intelligence tanpa Surveillance

Volunteer Intelligence tanpa Surveillance ditempatkan sebagai unit kerja riset karena memetakan kompetensi relawan berdasarkan data yang mereka berikan secara sadar untuk penugasan yang lebih baik. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UU 27/2022; OECD, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi skills inventory. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah volunteer matching, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.8. Rencana tindakan riset: Volunteer Intelligence tanpa Surveillance

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi volunteer matching dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan skills inventory.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur task fit, retention.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: profiling, coercion.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam volunteer intelligence tanpa surveillance, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah task fit, retention. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UU 27/2022; OECD, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan skills inventory. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi volunteer matching: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan

indikator task fit, retention dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa task fit, retention; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi profiling, coercion. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, volunteer matching memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UU 27/2022; OECD, 2024).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas volunteer matching, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

6.8 Evaluasi Dampak Program Masjid

Evaluasi Dampak Program Masjid ditempatkan sebagai unit kerja riset karena mengukur jangkauan, kualitas implementasi, hasil belajar, perubahan praktik, kepuasan, serta keberlanjutan. Fokus pada masjid mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Glasgow et al., 1999; Creswell & Plano Clark, 2018. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi RE-AIM + mixed methods. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah impact report, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 6.9. Rencana tindakan riset: Evaluasi Dampak Program Masjid

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
2. Desain	Susun spesifikasi impact report dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan RE-AIM + mixed methods.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur reach, adoption, effectiveness, maintenance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: vanity metrics.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

6.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam evaluasi dampak program masjid, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah reach, adoption, effectiveness, maintenance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Glasgow et al., 1999; Creswell & Plano Clark, 2018 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

6.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan RE-AIM + mixed methods. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi impact report: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator reach, adoption, effectiveness, maintenance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila

kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

6.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa reach, adoption, effectiveness, maintenance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi vanity metrics. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, impact report memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Glasgow et al., 1999; Creswell & Plano Clark, 2018).

Contoh penerapan riil: pada masjid, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas impact report, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 6

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjalankan diagnosis cepat masalah layanan, pembelajaran, komunikasi, akses, dan kebutuhan kelompok jamaah' sebagai contoh. Bukti dapat berasal dari survei singkat, wawancara, observasi, sedangkan produk atau intervensinya dapat berbentuk baseline kebutuhan jamaah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti response coverage, priority score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti survei bias, data berlebih perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengonsolidasikan SOP, jadwal, materi kajian, program sosial, panduan ibadah institusional, dan kontak layanan dalam basis pengetahuan' sebagai contoh. Bukti dapat berasal dari content inventory + RAG, sedangkan produk atau intervensinya dapat berbentuk knowledge hub. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti search success, freshness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dokumen kedaluwarsa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memberi jawaban tentang jadwal, lokasi, pendaftaran kegiatan, materi rujukan, dan FAQ administrasi dengan eskalasi manusia' sebagai contoh. Bukti dapat berasal dari service design + chatbot pilot, sedangkan produk atau intervensinya dapat berbentuk service assistant. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti resolution rate, escalation accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pertanyaan hukum agama masuk scope perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mentranskripsi, memberi metadata, ringkasan, tema, dan tautan sumber untuk membangun memori institusional' sebagai contoh. Bukti dapat berasal dari ASR + curation, sedangkan produk atau intervensinya dapat berbentuk digital sermon archive. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti searchability, transcript accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti copyright, misquote perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghubungkan program dakwah dengan data kebutuhan pendidikan, kemiskinan, kesehatan sosial, keluarga, dan relawan secara agregat' sebagai contoh. Bukti dapat berasal dari needs assessment + program evaluation, sedangkan produk atau intervensinya dapat berbentuk dashboard program sosial. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, referral, completion harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti stigma, sensitive data perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mendidik jamaah untuk memeriksa sumber, mengenali konten sintetis, menjaga privasi, dan menggunakan AI secara proporsional' sebagai contoh. Bukti dapat berasal dari pre-post workshop, sedangkan produk atau intervensinya dapat berbentuk modul literasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti knowledge gain, detection accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti fear-mongering atau overconfidence perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memetakan kompetensi relawan berdasarkan data yang mereka berikan secara sadar untuk penugasan yang lebih baik' sebagai contoh. Bukti dapat berasal dari skills inventory, sedangkan produk atau intervensinya dapat berbentuk volunteer matching. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti task fit, retention harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling, coercion perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengukur jangkauan, kualitas implementasi, hasil belajar, perubahan praktik, kepuasan, serta keberlanjutan' sebagai contoh. Bukti dapat berasal dari RE-AIM + mixed methods, sedangkan produk atau intervensinya dapat berbentuk impact report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, adoption, effectiveness, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti vanity metrics perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjalankan diagnosis cepat masalah layanan, pembelajaran, komunikasi, akses, dan kebutuhan kelompok jamaah' sebagai contoh. Bukti dapat berasal dari survei singkat, wawancara, observasi, sedangkan produk atau intervensinya dapat berbentuk baseline kebutuhan jamaah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti response coverage, priority score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti survei bias, data berlebih perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengonsolidasikan SOP, jadwal, materi kajian, program sosial, panduan ibadah institusional, dan kontak layanan dalam basis pengetahuan' sebagai contoh. Bukti dapat berasal dari content inventory + RAG, sedangkan produk atau intervensinya dapat berbentuk knowledge hub. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti search success, freshness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dokumen kedaluwarsa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memberi jawaban tentang jadwal, lokasi, pendaftaran kegiatan, materi rujukan, dan FAQ administrasi dengan eskalasi manusia' sebagai contoh. Bukti dapat berasal dari service design + chatbot pilot, sedangkan produk atau intervensinya dapat berbentuk service assistant. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti resolution rate, escalation accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pertanyaan hukum agama masuk scope perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mentranskripsi, memberi metadata, ringkasan, tema, dan tautan sumber untuk membangun memori institusional' sebagai contoh. Bukti dapat berasal dari ASR + curation, sedangkan produk atau intervensinya dapat berbentuk digital sermon archive. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti searchability, transcript accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti copyright, misquote perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks komunitas, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghubungkan program dakwah dengan data kebutuhan pendidikan, kemiskinan, kesehatan sosial, keluarga, dan relawan secara agregat' sebagai contoh. Bukti dapat berasal dari needs assessment + program evaluation, sedangkan produk atau intervensinya dapat berbentuk dashboard program sosial. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, referral, completion harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti stigma, sensitive data perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mendidik jamaah untuk memeriksa sumber, mengenali konten sintetis, menjaga privasi, dan menggunakan AI secara proporsional' sebagai contoh. Bukti dapat berasal dari pre-post workshop, sedangkan produk atau intervensinya dapat berbentuk modul literasi. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti knowledge gain, detection accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti fear-mongering atau overconfidence perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memetakan kompetensi relawan berdasarkan data yang mereka berikan secara sadar untuk penugasan yang lebih baik' sebagai contoh. Bukti dapat berasal dari skills inventory, sedangkan produk atau intervensinya dapat berbentuk volunteer matching. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti task fit, retention harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti profiling, coercion perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengukur jangkauan, kualitas implementasi, hasil belajar, perubahan praktik, kepuasan, serta keberlanjutan' sebagai contoh. Bukti dapat berasal dari RE-AIM + mixed methods, sedangkan produk atau intervensinya dapat berbentuk impact report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, adoption, effectiveness, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti vanity metrics perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjalankan diagnosis cepat masalah layanan, pembelajaran, komunikasi, akses, dan kebutuhan kelompok jamaah' sebagai contoh. Bukti dapat berasal dari survei singkat, wawancara, observasi, sedangkan produk atau intervensinya dapat berbentuk baseline kebutuhan jamaah. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti response coverage, priority score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti survei bias, data berlebih perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengonsolidasikan SOP, jadwal, materi kajian, program sosial, panduan ibadah institusional, dan kontak layanan dalam basis pengetahuan' sebagai contoh. Bukti dapat berasal dari content inventory + RAG, sedangkan produk atau intervensinya dapat berbentuk knowledge hub. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti search success, freshness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dokumen kedaluwarsa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memberi jawaban tentang jadwal, lokasi, pendaftaran kegiatan, materi rujukan, dan FAQ administrasi dengan eskalasi manusia' sebagai contoh. Bukti dapat berasal dari service design + chatbot pilot, sedangkan produk atau intervensinya dapat berbentuk service assistant. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti resolution rate, escalation accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pertanyaan hukum agama masuk scope perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mentranskripsi, memberi metadata, ringkasan, tema, dan tautan sumber untuk membangun memori institusional' sebagai contoh. Bukti dapat berasal dari ASR + curation, sedangkan produk atau intervensinya dapat berbentuk digital sermon archive. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti searchability, transcript accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti copyright, misquote perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

SOAL PENGEMBANGAN BAB 6

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 6 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 6, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 7 — IMPLEMENTASI RIIL DI SEKOLAH DAN KAMPUS

Bab 7 membangun kompetensi untuk implementasi riil di sekolah dan kampus. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 7



Gambar 7.1. Kerangka operasional Implementasi Riil Di Sekolah Dan Kampus

Tabel 7.1. Matriks kerja utama Bab 7

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan
Desain	Bagaimana solusi bekerja	Spesifikasi + korpus	Prototipe	Uji internal

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
	dan di mana batasnya?			
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

7.1 AI Literacy dalam Pendidikan Agama

AI Literacy dalam Pendidikan Agama ditempatkan sebagai unit kerja riset karena mengajarkan konsep AI, keterbatasan, verifikasi, etika, dan penggunaan bertanggung jawab dalam konteks pembelajaran agama. Fokus pada sekolah/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Miao & Holmes, 2023; Kasneci et al., 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi curriculum design + pre-post. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah modul AI literacy, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.2. Rencana tindakan riset: AI Literacy dalam Pendidikan Agama

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi modul AI literacy dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan curriculum design + pre-post.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur knowledge, critical evaluation.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: ketergantungan AI.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam ai literacy dalam pendidikan agama, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah knowledge, critical evaluation. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Miao & Holmes, 2023; Kasneci et al., 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan curriculum design + pre-post. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi modul AI literacy: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim

mengumpulkan indikator knowledge, critical evaluation dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa knowledge, critical evaluation; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi ketergantungan AI. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, modul AI literacy memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna

menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Miao & Holmes, 2023; Kasneci et al., 2023).

Contoh penerapan riil: pada sekolah/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas modul AI literacy, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.2 Tutor Pendamping Guru Berbasis Korpus

Tutor Pendamping Guru Berbasis Korpus ditempatkan sebagai unit kerja riset karena membuat tutor yang menjelaskan materi berdasarkan bahan ajar institusi dan selalu memberi sumber serta pertanyaan reflektif. Fokus pada sekolah mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2023; Lewis et al., 2020. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi RAG + teacher review. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah tutor pendamping, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.3. Rencana tindakan riset: Tutor Pendamping Guru Berbasis Korpus

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah; prioritaskan	Ringkasan baseline,	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	satu masalah.	consent, catatan sampling	
2. Desain	Susun spesifikasi tutor pendamping dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan RAG + teacher review.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur learning gain, source fidelity.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: teacher displacement narrative.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam tutor pendamping guru berbasis korpus, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah learning gain, source fidelity. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2023; Lewis et al., 2020 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan RAG + teacher review. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi tutor pendamping: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator learning gain, source fidelity dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa learning gain, source fidelity; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi teacher displacement narrative. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, tutor pendamping memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2023; Lewis et al., 2020).

Contoh penerapan riil: pada sekolah, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas tutor pendamping, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.3 Microlearning dan Retrieval Practice

Microlearning dan Retrieval Practice ditempatkan sebagai unit kerja riset karena menghasilkan kartu belajar, pertanyaan, dan feedback dari materi yang telah disahkan untuk mendukung pengulangan terjadwal. Fokus pada sekolah mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Kasneci et al., 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi quasi-experiment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah microlearning pack, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.4. Rencana tindakan riset: Microlearning dan Retrieval Practice

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi microlearning pack dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
3. Prototipe	Bangun versi minimum menggunakan quasi-experiment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur retention, completion.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: soal ambigu, answer leakage.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam microlearning dan retrieval practice, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah retention, completion. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Kasneci et al., 2023 memberi kerangka awal, sementara konteks organisasi

menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan quasi-experiment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi microlearning pack: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator retention, completion dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa retention, completion; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi soal ambigu, answer leakage. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, microlearning pack memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Kasneci et al., 2023).

Contoh penerapan riil: pada sekolah, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas microlearning pack, dokumentasi sumber, instrumen

evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.4 Laboratorium Dakwah Digital Mahasiswa

Laboratorium Dakwah Digital Mahasiswa ditempatkan sebagai unit kerja riset karena membentuk studio riset tempat mahasiswa melakukan needs assessment, prototyping, uji pengguna, dan evaluasi etika. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain McKenney & Reeves, 2019. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi project-based EDR. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah portfolio proyek, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.5. Rencana tindakan riset: Laboratorium Dakwah Digital Mahasiswa

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi portfolio proyek dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan project-based EDR.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur research quality, usability.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
5. Keputusan	Tinjau risiko: proyek mengejar tren.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam laboratorium dakwah digital mahasiswa, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah research quality, usability. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. McKenney & Reeves, 2019 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan project-based EDR. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi portfolio proyek: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator research quality, usability dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa research quality, usability; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi proyek mengejar tren. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, portfolio proyek memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (McKenney & Reeves, 2019).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas portfolio proyek, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.5 Digital Ethnography Komunitas Kampus

Digital Ethnography Komunitas Kampus ditempatkan sebagai unit kerja riset karena mengamati praktik, wacana, dan kebutuhan beragama di ruang digital

dengan protokol etik. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Campbell & Tsuria, 2022; AoIR, 2019. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi digital ethnography. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah insight report, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.6. Rencana tindakan riset: Digital Ethnography Komunitas Kampus

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi insight report dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan digital ethnography.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur thematic saturation, ethics compliance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: privacy intrusion.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam digital ethnography komunitas kampus, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah thematic saturation, ethics compliance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Campbell & Tsuria, 2022; AoIR, 2019 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan digital ethnography. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi insight report: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator thematic saturation, ethics compliance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa thematic saturation, ethics compliance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi privacy intrusion. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, insight report memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Campbell & Tsuria, 2022; AoIR, 2019).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas insight report, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.6 Eksperimen Pesan Dakwah secara Etis

Eksperimen Pesan Dakwah secara Etis ditempatkan sebagai unit kerja riset karena membandingkan variasi format pesan untuk pemahaman dan keterlibatan tanpa manipulasi emosional atau deception yang tidak perlu. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain OECD, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta

batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi A/B test/field experiment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah evidence-based message design, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.7. Rencana tindakan riset: Eksperimen Pesan Dakwah secara Etis

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi evidence-based message design dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan A/B test/field experiment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur comprehension, voluntary engagement.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: dark patterns, deceptive experimentation.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam eksperimen pesan dakwah secara etis, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah comprehension, voluntary engagement. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. OECD, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan A/B test/field experiment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi evidence-based message design: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara

internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator comprehension, voluntary engagement dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa comprehension, voluntary engagement; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi dark patterns, deceptive experimentation. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, evidence-based message design memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik

sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (OECD, 2024).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas evidence-based message design, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.7 Penilaian Autentik di Era GenAI

Penilaian Autentik di Era GenAI ditempatkan sebagai unit kerja riset karena menggeser asesmen dari produk akhir semata ke proses, oral defense, provenance, log revisi, dan refleksi. Fokus pada sekolah/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2023; Stanford HAI, 2026. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi assessment redesign. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah authentic assessment, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.8. Rencana tindakan riset: Penilaian Autentik di Era GenAI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada sekolah/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi authentic assessment dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan assessment redesign.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur process evidence, mastery.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: false accusation, surveillance.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam penilaian autentik di era genai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah process evidence, mastery. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada sekolah/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2023; Stanford HAI, 2026 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan assessment redesign. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi authentic assessment: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada sekolah/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim

mengumpulkan indikator process evidence, mastery dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa process evidence, mastery; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi false accusation, surveillance. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, authentic assessment memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks sekolah/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2023; Stanford HAI, 2026).

Contoh penerapan riil: pada sekolah/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas authentic assessment, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

7.8 Capstone Riset-Pengembangan Dakwah AI

Capstone Riset-Pengembangan Dakwah AI ditempatkan sebagai unit kerja riset karena mengintegrasikan masalah nyata, literature review, desain artefak, ethics review, pilot, evaluasi, dan publikasi. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Peffers et al., 2007; McKenney & Reeves, 2019. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi DSRM/EDR/mixed methods. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah capstone prototype + report, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 7.9. Rencana tindakan riset: Capstone Riset-Pengembangan Dakwah AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		sampling	
2. Desain	Susun spesifikasi capstone prototype + report dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan DSRM/EDR/mixed methods.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur utility, rigor, impact.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: scope terlalu besar.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

7.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam capstone riset-pengembangan dakwah ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah utility, rigor, impact. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Peffers et al., 2007; McKenney & Reeves, 2019 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

7.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan DSRM/EDR/mixed methods. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi capstone prototype + report: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator utility, rigor, impact dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

7.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa utility, rigor, impact; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi scope terlalu besar. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, capstone prototype + report memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Peffer et al., 2007; McKenney & Reeves, 2019).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas capstone prototype + report, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 7

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengajarkan konsep AI, keterbatasan, verifikasi, etika, dan penggunaan bertanggung jawab dalam konteks pembelajaran agama' sebagai contoh. Bukti dapat berasal dari curriculum design + pre-post, sedangkan produk atau intervensinya dapat berbentuk modul AI literacy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti knowledge, critical evaluation harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti ketergantungan AI perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat tutor yang menjelaskan materi berdasarkan bahan ajar institusi dan selalu memberi sumber serta pertanyaan reflektif' sebagai contoh. Bukti dapat berasal dari RAG + teacher review, sedangkan produk atau intervensinya dapat berbentuk tutor pendamping. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, source fidelity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti teacher displacement narrative perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghasilkan kartu belajar, pertanyaan, dan feedback dari materi yang telah disahkan untuk mendukung pengulangan terjadwal' sebagai contoh. Bukti dapat berasal dari quasi-experiment, sedangkan produk atau intervensinya dapat berbentuk microlearning pack. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retention, completion harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti soal ambigu, answer leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membentuk studio riset tempat mahasiswa melakukan needs assessment, prototyping, uji pengguna, dan evaluasi etika' sebagai contoh. Bukti dapat berasal dari project-based EDR, sedangkan produk atau intervensinya dapat berbentuk portfolio proyek. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti research quality, usability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti proyek mengejar tren perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengamati praktik, wacana, dan kebutuhan beragama di ruang digital dengan protokol etik' sebagai contoh. Bukti dapat berasal dari digital ethnography, sedangkan produk atau intervensinya dapat berbentuk insight report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti thematic saturation, ethics compliance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti privacy intrusion perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membandingkan variasi format pesan untuk pemahaman dan keterlibatan tanpa manipulasi emosional atau deception yang tidak perlu' sebagai contoh. Bukti dapat berasal dari A/B test/field experiment, sedangkan produk atau intervensinya dapat berbentuk evidence-based message design. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti comprehension, voluntary engagement harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dark patterns, deceptive experimentation perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggeser asesmen dari produk akhir semata ke proses, oral defense, provenance, log revisi, dan refleksi' sebagai contoh. Bukti dapat berasal dari assessment redesign, sedangkan produk atau intervensinya dapat berbentuk authentic assessment. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti process evidence, mastery harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti false accusation, surveillance perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengintegrasikan masalah nyata, literature review, desain artefak, ethics review, pilot, evaluasi, dan publikasi' sebagai contoh. Bukti dapat berasal dari DSRM/EDR/mixed methods, sedangkan produk atau intervensinya dapat berbentuk capstone prototype + report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti utility, rigor, impact harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti scope terlalu besar perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengajarkan konsep AI, keterbatasan, verifikasi, etika, dan penggunaan bertanggung jawab dalam konteks pembelajaran agama' sebagai contoh. Bukti dapat berasal dari curriculum design + pre-post, sedangkan produk atau intervensinya dapat berbentuk modul AI literacy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti knowledge, critical evaluation harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti ketergantungan AI perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat tutor yang menjelaskan materi berdasarkan bahan ajar institusi dan selalu memberi sumber serta pertanyaan reflektif' sebagai contoh. Bukti dapat berasal dari RAG + teacher review, sedangkan produk atau intervensinya dapat berbentuk tutor pendamping. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, source fidelity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti teacher displacement narrative perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghasilkan kartu belajar, pertanyaan, dan feedback dari materi yang telah disahkan untuk mendukung pengulangan terjadwal' sebagai contoh. Bukti dapat berasal dari quasi-experiment, sedangkan produk atau intervensinya dapat berbentuk microlearning pack. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retention, completion harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti soal ambigu, answer leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membentuk studio riset tempat mahasiswa melakukan needs assessment, prototyping, uji pengguna, dan evaluasi etika' sebagai contoh. Bukti dapat berasal dari project-based EDR, sedangkan produk atau intervensinya dapat berbentuk portfolio proyek. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti research quality, usability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti proyek mengejar tren perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengamati praktik, wacana, dan kebutuhan beragama di ruang digital dengan protokol etik' sebagai contoh. Bukti dapat berasal dari digital ethnography, sedangkan produk atau intervensinya dapat berbentuk insight report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti thematic saturation, ethics compliance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti privacy intrusion perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membandingkan variasi format pesan untuk pemahaman dan keterlibatan tanpa manipulasi emosional atau deception yang tidak perlu' sebagai contoh. Bukti dapat berasal dari A/B test/field experiment, sedangkan produk atau intervensinya dapat berbentuk evidence-based message design. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti comprehension, voluntary engagement harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti dark patterns, deceptive experimentation perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggeser asesmen dari produk akhir semata ke proses, oral defense, provenance, log revisi, dan refleksi' sebagai contoh. Bukti dapat berasal dari assessment redesign, sedangkan produk atau intervensinya dapat berbentuk authentic assessment. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti process evidence, mastery harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti false accusation, surveillance perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengintegrasikan masalah nyata, literature review, desain artefak, ethics review, pilot, evaluasi, dan publikasi' sebagai contoh. Bukti dapat berasal dari DSRM/EDR/mixed methods, sedangkan produk atau intervensinya dapat berbentuk capstone prototype + report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti utility, rigor, impact harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti scope terlalu besar perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengajarkan konsep AI, keterbatasan, verifikasi, etika, dan penggunaan bertanggung jawab dalam konteks pembelajaran agama' sebagai contoh. Bukti dapat berasal dari curriculum design + pre-post, sedangkan produk atau intervensinya dapat berbentuk modul AI literacy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti knowledge, critical evaluation harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti ketergantungan AI perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat tutor yang menjelaskan materi berdasarkan bahan ajar institusi dan selalu memberi sumber serta pertanyaan reflektif' sebagai contoh. Bukti dapat berasal dari RAG + teacher review, sedangkan produk atau intervensinya dapat berbentuk tutor pendamping. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti learning gain, source fidelity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti teacher displacement narrative perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghasilkan kartu belajar, pertanyaan, dan feedback dari materi yang telah disahkan untuk mendukung pengulangan terjadwal' sebagai contoh. Bukti dapat berasal dari quasi-experiment, sedangkan produk atau intervensinya dapat berbentuk microlearning pack. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti retention, completion harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti soal ambigu, answer leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membentuk studio riset tempat mahasiswa melakukan needs assessment, prototyping, uji pengguna, dan evaluasi etika' sebagai contoh. Bukti dapat berasal dari project-based EDR, sedangkan produk atau intervensinya dapat berbentuk portfolio proyek. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti research quality, usability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti proyek mengejar tren perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

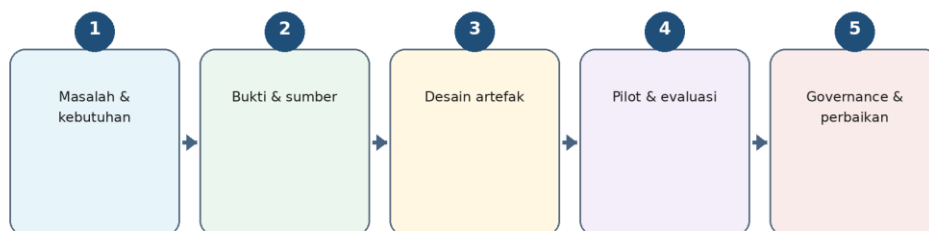
SOAL PENGEMBANGAN BAB 7

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 7 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 7, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 8 — IMPLEMENTASI RIIL PADA LEMBAGA DAKWAH

Bab 8 membangun kompetensi untuk implementasi riil pada lembaga dakwah. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 8



Gambar 8.1. Kerangka operasional Implementasi Riil Pada Lembaga Dakwah

Tabel 8.1. Matriks kerja utama Bab 8

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan
Desain	Bagaimana solusi bekerja	Spesifikasi + korpus	Prototipe	Uji internal

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
	dan di mana batasnya?			
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

8.1 Strategi Portofolio AI Lembaga

Strategi Portofolio AI Lembaga ditempatkan sebagai unit kerja riset karena memilih kasus penggunaan berdasarkan nilai, risiko, biaya, kesiapan data, dan kapasitas manusia. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain ISO/IEC 42001, 2023; NIST, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi portfolio scoring. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah AI portfolio roadmap, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.2. Rencana tindakan riset: Strategi Portofolio AI Lembaga

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei	Ringkasan	Masalah spesifik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	pada lembaga; prioritaskan satu masalah.	baseline, consent, catatan sampling	dan terukur
2. Desain	Susun spesifikasi AI portfolio roadmap dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan portfolio scoring.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur benefit-risk score.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: AI theater dan vendor lock-in.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam strategi portofolio ai lembaga, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah benefit-risk score. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. ISO/IEC 42001, 2023; NIST, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan portfolio scoring. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi AI portfolio roadmap: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator benefit-risk score dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa benefit-risk score; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi AI theater dan vendor lock-in. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, AI portfolio roadmap memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (ISO/IEC 42001, 2023; NIST, 2023).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas AI portfolio roadmap, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.2 AI Center of Excellence Skala Lembaga

AI Center of Excellence Skala Lembaga ditempatkan sebagai unit kerja riset karena membangun tim lintas fungsi untuk standar, eksperimen, pendidikan, review risiko, dan dokumentasi. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain ISO/IEC 42001, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi organizational design. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah AI CoE, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.3. Rencana tindakan riset: AI Center of Excellence Skala Lembaga

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi AI CoE dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
3. Prototipe	Bangun versi minimum menggunakan organizational design.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur policy adoption, review SLA.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: sentralisasi berlebihan.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam ai center of excellence skala lembaga, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah policy adoption, review SLA. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. ISO/IEC 42001, 2023 memberi kerangka awal, sementara konteks

organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan organizational design. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi AI CoE: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator policy adoption, review SLA dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa policy adoption, review SLA; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi sentralisasi berlebihan. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, AI CoE memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (ISO/IEC 42001, 2023).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas AI CoE, dokumentasi sumber, instrumen evaluasi, risk

register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.3 SOP Pengadaan dan Evaluasi Vendor AI

SOP Pengadaan dan Evaluasi Vendor AI ditempatkan sebagai unit kerja riset karena menilai keamanan, privasi, data processing, transparansi, portabilitas, biaya, dan mekanisme penghentian layanan. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024; UU 27/2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi vendor assessment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah procurement checklist, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.4. Rencana tindakan riset: SOP Pengadaan dan Evaluasi Vendor AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi procurement checklist dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan vendor assessment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur due diligence completeness.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
5. Keputusan	Tinjau risiko: data residency, hidden retention.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam sop pengadaan dan evaluasi vendor ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah due diligence completeness. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024; UU 27/2022 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan vendor assessment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi procurement checklist: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator due diligence completeness dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa due diligence completeness; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi data residency, hidden retention. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, procurement checklist memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024; UU 27/2022).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas procurement checklist, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.4 Dashboard Dakwah Berbasis Outcome

Dashboard Dakwah Berbasis Outcome ditempatkan sebagai unit kerja riset karena mengintegrasikan data program secara agregat untuk melihat reach, learning,

service quality, dan follow-up tanpa mengubah dakwah menjadi sekadar metrik. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain RE-AIM; Rogers, 2003. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi measurement design. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah impact dashboard, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.5. Rencana tindakan riset: Dashboard Dakwah Berbasis Outcome

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi impact dashboard dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan measurement design.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur outcome coverage, data quality.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: metric gaming.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam dashboard dakwah berbasis outcome, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah outcome coverage, data quality. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. RE-AIM; Rogers, 2003 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan measurement design. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi impact dashboard: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator outcome coverage, data quality dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa outcome coverage, data quality; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi metric gaming. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, impact dashboard memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (RE-AIM; Rogers, 2003).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas impact dashboard, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.5 Multibahasa dan Skala Distribusi Konten

Multibahasa dan Skala Distribusi Konten ditempatkan sebagai unit kerja riset karena menerjemahkan dan mengadaptasi konten dengan review penutur/ahli serta memelihara versi yang terlacak. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta

batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi translation QA. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah multilingual content pipeline, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.6. Rencana tindakan riset: Multibahasa dan Skala Distribusi Konten

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi multilingual content pipeline dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan translation QA.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur terminology accuracy.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: context loss.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam multibahasa dan skala distribusi konten, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah terminology accuracy. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan translation QA. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi multilingual content pipeline: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara

internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator terminology accuracy dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa terminology accuracy; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi context loss. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, multilingual content pipeline memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden,

perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas multilingual content pipeline, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.6 Knowledge Governance dan Editorial Board

Knowledge Governance dan Editorial Board ditempatkan sebagai unit kerja riset karena menetapkan siapa yang berhak menyetujui sumber, revisi, taksonomi, jawaban benchmark, dan penanganan sengketa. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Bunt, 2018; ISO/IEC 42001, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi governance charter. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah editorial governance, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.7. Rencana tindakan riset: Knowledge Governance dan Editorial Board

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi editorial governance dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan governance charter.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur decision traceability.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: authority ambiguity.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam knowledge governance dan editorial board, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah decision traceability. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Bunt, 2018; ISO/IEC 42001, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan governance charter. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi editorial governance: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan

indikator decision traceability dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa decision traceability; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi authority ambiguity. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, editorial governance memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Bunt, 2018; ISO/IEC 42001, 2023).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas editorial governance, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.7 Incident Response untuk Kesalahan AI

Incident Response untuk Kesalahan AI ditempatkan sebagai unit kerja riset karena membuat jalur pelaporan, triase, containment, koreksi publik, root-cause analysis, dan pencegahan pengulangan. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024; Stanford HAI, 2026. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi incident management. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah AI incident playbook, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.8. Rencana tindakan riset: Incident Response untuk Kesalahan AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
2. Desain	Susun spesifikasi AI incident playbook dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan incident management.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur time-to-detect, time-to-correct.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: menutup-nutupi insiden.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam incident response untuk kesalahan ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah time-to-detect, time-to-correct. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024; Stanford HAI, 2026 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan incident management. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi AI incident playbook: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator time-to-detect, time-to-correct dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena

R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa time-to-detect, time-to-correct; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi menutup-nutupi insiden. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, AI incident playbook memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024; Stanford HAI, 2026).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara,

analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas AI incident playbook, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

8.8 Model Kematangan Dakwah AI

Model Kematangan Dakwah AI ditempatkan sebagai unit kerja riset karena menilai kematangan dari ad hoc, managed, governed, measured, hingga continuously improved. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain ISO/IEC 42001, 2023; OECD, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi maturity assessment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah maturity roadmap, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 8.9. Rencana tindakan riset: Model Kematangan Dakwah AI

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi maturity roadmap dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan maturity assessment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur capability	Dataset pilot,	Data cukup untuk

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	score.	log insiden, wawancara	keputusan
5. Keputusan	Tinjau risiko: self-rating bias.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

8.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam model kematangan dakwah ai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah capability score. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. ISO/IEC 42001, 2023; OECD, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

8.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan maturity assessment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi maturity roadmap: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator capability score dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

8.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa capability score; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi self-rating bias. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, maturity roadmap memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (ISO/IEC 42001, 2023; OECD, 2024).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas maturity roadmap, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 8

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memilih kasus penggunaan berdasarkan nilai, risiko, biaya, kesiapan data, dan kapasitas manusia' sebagai contoh. Bukti dapat berasal dari portfolio scoring, sedangkan produk atau intervensinya dapat berbentuk AI portfolio roadmap. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti benefit-risk score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI theater dan vendor lock-in perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun tim lintas fungsi untuk standar, eksperimen, pendidikan, review risiko, dan dokumentasi' sebagai contoh. Bukti dapat berasal dari organizational design, sedangkan produk atau intervensinya dapat berbentuk AI CoE. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti policy adoption, review SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti sentralisasi berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai keamanan, privasi, data processing, transparansi, portabilitas, biaya, dan mekanisme penghentian layanan' sebagai contoh. Bukti dapat berasal dari vendor assessment, sedangkan produk atau intervensinya dapat berbentuk procurement checklist. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti due diligence completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data residency, hidden retention perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengintegrasikan data program secara agregat untuk melihat reach, learning, service quality, dan follow-up tanpa mengubah dakwah menjadi sekadar metrik' sebagai contoh. Bukti dapat berasal dari measurement design, sedangkan produk atau intervensinya dapat berbentuk impact dashboard. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti outcome coverage, data quality harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti metric gaming perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerjemahkan dan mengadaptasi konten dengan review penutur/ahli serta memelihara versi yang terlacak' sebagai contoh. Bukti dapat berasal dari translation QA, sedangkan produk atau intervensinya dapat berbentuk multilingual content pipeline. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti terminology accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti context loss perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menetapkan siapa yang berhak menyetujui sumber, revisi, taksonomi, jawaban benchmark, dan penanganan sengketa' sebagai contoh. Bukti dapat berasal dari governance charter, sedangkan produk atau intervensinya dapat berbentuk editorial governance. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti decision traceability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti authority ambiguity perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat jalur pelaporan, triase, containment, koreksi publik, root-cause analysis, dan pencegahan pengulangan' sebagai contoh. Bukti dapat berasal dari incident management, sedangkan produk atau intervensinya dapat berbentuk AI incident playbook. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti time-to-detect, time-to-correct harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti menutup-nutupi insiden perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai kematangan dari ad hoc, managed, governed, measured, hingga continuously improved' sebagai contoh. Bukti dapat berasal dari maturity assessment, sedangkan produk atau intervensinya dapat berbentuk maturity roadmap. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti capability score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti self-rating bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memilih kasus penggunaan berdasarkan nilai, risiko, biaya, kesiapan data, dan kapasitas manusia' sebagai contoh. Bukti dapat berasal dari portfolio scoring, sedangkan produk atau intervensinya dapat berbentuk AI portfolio roadmap. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti benefit-risk score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI theater dan vendor lock-in perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun tim lintas fungsi untuk standar, eksperimen, pendidikan, review risiko, dan dokumentasi' sebagai contoh. Bukti dapat berasal dari organizational design, sedangkan produk atau intervensinya dapat berbentuk AI CoE. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti policy adoption, review SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti sentralisasi berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai keamanan, privasi, data processing, transparansi, portabilitas, biaya, dan mekanisme penghentian layanan' sebagai contoh. Bukti dapat berasal dari vendor assessment, sedangkan produk atau intervensinya dapat berbentuk procurement checklist. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti due diligence completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data residency, hidden retention perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengintegrasikan data program secara agregat untuk melihat reach, learning, service quality, dan follow-up tanpa mengubah dakwah menjadi sekadar metrik' sebagai contoh. Bukti dapat berasal dari measurement design, sedangkan produk atau intervensinya dapat berbentuk impact dashboard. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti outcome coverage, data quality harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti metric gaming perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerjemahkan dan mengadaptasi konten dengan review penutur/ahli serta memelihara versi yang terlacak' sebagai contoh. Bukti dapat berasal dari translation QA, sedangkan produk atau intervensinya dapat berbentuk multilingual content pipeline. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti terminology accuracy harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti context loss perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menetapkan siapa yang berhak menyetujui sumber, revisi, taksonomi, jawaban benchmark, dan penanganan sengketa' sebagai contoh. Bukti dapat berasal dari governance charter, sedangkan produk atau intervensinya dapat berbentuk editorial governance. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti decision traceability harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti authority ambiguity perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membuat jalur pelaporan, triase, containment, koreksi publik, root-cause analysis, dan pencegahan pengulangan' sebagai contoh. Bukti dapat berasal dari incident management, sedangkan produk atau intervensinya dapat berbentuk AI incident playbook. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti time-to-detect, time-to-correct harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti menutup-nutupi insiden perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai kematangan dari ad hoc, managed, governed, measured, hingga continuously improved' sebagai contoh. Bukti dapat berasal dari maturity assessment, sedangkan produk atau intervensinya dapat berbentuk maturity roadmap. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti capability score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti self-rating bias perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memilih kasus penggunaan berdasarkan nilai, risiko, biaya, kesiapan data, dan kapasitas manusia' sebagai contoh. Bukti dapat berasal dari portfolio scoring, sedangkan produk atau intervensinya dapat berbentuk AI portfolio roadmap. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti benefit-risk score harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti AI theater dan vendor lock-in perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'membangun tim lintas fungsi untuk standar, eksperimen, pendidikan, review risiko, dan dokumentasi' sebagai contoh. Bukti dapat berasal dari organizational design, sedangkan produk atau intervensinya dapat berbentuk AI CoE. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti policy adoption, review SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti sentralisasi berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai keamanan, privasi, data processing, transparansi, portabilitas, biaya, dan mekanisme penghentian layanan' sebagai contoh. Bukti dapat berasal dari vendor assessment, sedangkan produk atau intervensinya dapat berbentuk procurement checklist. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti due diligence completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data residency, hidden retention perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengintegrasikan data program secara agregat untuk melihat reach, learning, service quality, dan follow-up tanpa mengubah dakwah menjadi sekadar metrik' sebagai contoh. Bukti dapat berasal dari measurement design, sedangkan produk atau intervensinya dapat berbentuk impact dashboard. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti outcome coverage, data quality harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti metric gaming perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

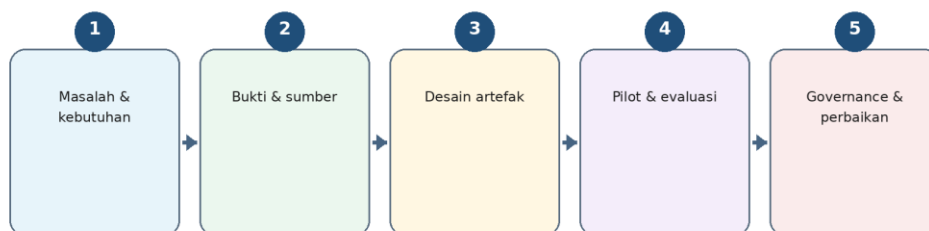
SOAL PENGEMBANGAN BAB 8

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 8 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 8, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 9 — ETIKA, HUKUM, KEAMANAN, DAN OTORITAS

Bab 9 membangun kompetensi untuk etika, hukum, keamanan, dan otoritas. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 9



Gambar 9.1. Kerangka operasional Etika, Hukum, Keamanan, Dan Otoritas

Tabel 9.1. Matriks kerja utama Bab 9

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan
Desain	Bagaimana solusi bekerja	Spesifikasi + korpus	Prototipe	Uji internal

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
	dan di mana batasnya?			
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

9.1 Maqasid, Masalah, dan Prinsip Do No Harm

Maqasid, Masalah, dan Prinsip Do No Harm ditempatkan sebagai unit kerja riset karena menggunakan orientasi kemaslahatan dan pencegahan mudarat untuk menilai proporsionalitas penggunaan AI. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021; Winter, 2026. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi ethical impact assessment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah ethical decision memo, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.2. Rencana tindakan riset: Maqasid, Masalah, dan Prinsip Do No Harm

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi ethical decision memo dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan ethical impact assessment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur benefit-harm balance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: techno-solutionism.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam maqasid, masalah, dan prinsip do no harm, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah benefit-harm balance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus

dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021; Winter, 2026 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan ethical impact assessment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi ethical decision memo: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator benefit-harm balance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa benefit-harm balance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi techno-solutionism. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, ethical decision memo memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021; Winter, 2026).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas ethical decision memo, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.2 Human Dignity, Fairness, dan Non-Diskriminasi

Human Dignity, Fairness, dan Non-Diskriminasi ditempatkan sebagai unit kerja riset karena menguji apakah sistem memperlakukan kelompok secara adil dan tidak memperburuk ketimpangan akses atau representasi. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UNESCO, 2021; OECD, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi fairness audit. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah fairness report, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.3. Rencana tindakan riset: Human Dignity, Fairness, dan Non-Diskriminasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei	Ringkasan	Masalah spesifik

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	pada lembaga; prioritaskan satu masalah.	baseline, consent, catatan sampling	dan terukur
2. Desain	Susun spesifikasi fairness report dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan fairness audit.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur error parity, access equity.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: bias data dan bahasa.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam human dignity, fairness, dan non-diskriminasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah error parity, access equity. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan

meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UNESCO, 2021; OECD, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan fairness audit. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi fairness report: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator error parity, access equity dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa error parity, access equity; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi bias data dan bahasa. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, fairness report memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UNESCO, 2021; OECD, 2024).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas fairness report, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.3 Pelindungan Data Pribadi dan Consent

Pelindungan Data Pribadi dan Consent ditempatkan sebagai unit kerja riset karena menerapkan tujuan spesifik, minimisasi data, keamanan, retention, hak subjek, serta transparansi pemrosesan. Fokus pada Indonesia mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain UU 27/2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi privacy impact assessment. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah privacy notice + register, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.4. Rencana tindakan riset: Pelindungan Data Pribadi dan Consent

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada Indonesia; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi privacy notice + register dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
3. Prototipe	Bangun versi minimum menggunakan privacy impact assessment.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur compliance, deletion SLA.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: unauthorized processing.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam perlindungan data pribadi dan consent, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah compliance, deletion SLA. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada Indonesia, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. UU 27/2022 memberi kerangka awal, sementara konteks organisasi

menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan privacy impact assessment. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi privacy notice + register: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada Indonesia, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator compliance, deletion SLA dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa compliance, deletion SLA; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi unauthorized processing. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, privacy notice + register memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks Indonesia, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (UU 27/2022).

Contoh penerapan riil: pada Indonesia, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas privacy notice + register, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut.

Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.4 Transparansi, Explainability, dan Provenance

Transparansi, Explainability, dan Provenance ditempatkan sebagai unit kerja riset karena menjelaskan fungsi, batas, sumber, versi model, penggunaan AI, dan siapa manusia penanggung jawab. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024; ISO/IEC 42001, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi documentation audit. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah model/system card, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.5. Rencana tindakan riset: Transparansi, Explainability, dan Provenance

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi model/system card dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan documentation audit.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur documentation	Dataset pilot, log insiden,	Data cukup untuk keputusan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	completeness.	wawancara	
5. Keputusan	Tinjau risiko: false transparency.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam transparansi, explainability, dan provenance, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah documentation completeness. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024; ISO/IEC 42001, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan documentation audit. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi model/system card: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator documentation completeness dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa documentation completeness; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi false transparency. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, model/system card memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024; ISO/IEC 42001, 2023).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas model/system card, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.5 Hallucination dan Reliability Engineering

Hallucination dan Reliability Engineering ditempatkan sebagai unit kerja riset karena mendeteksi keluaran tidak berdasar melalui retrieval grounding,

confidence/abstention, test set, adversarial testing, dan review manusia. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Ji et al., 2023; NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi benchmark + red teaming. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah reliability protocol, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.6. Rencana tindakan riset: Hallucination dan Reliability Engineering

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi reliability protocol dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan benchmark + red teaming.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur groundedness, hallucination rate.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: silent error.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam hallucination dan reliability engineering, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah groundedness, hallucination rate. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Ji et al., 2023; NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan benchmark + red teaming. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi reliability protocol: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator groundedness, hallucination rate dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa groundedness, hallucination rate; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi silent error. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, reliability protocol memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Ji et al., 2023; NIST, 2024).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas reliability protocol, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.6 Deepfake, Sintesis Suara, dan Identitas Dai

Deepfake, Sintesis Suara, dan Identitas Dai ditempatkan sebagai unit kerja riset karena mencegah pemalsuan identitas dan penggunaan suara/wajah tokoh tanpa izin serta menetapkan pelabelan konten sintesis. Fokus pada media mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain EU AI Act, 2024; NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta

batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi provenance and consent policy. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah synthetic media policy, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.7. Rencana tindakan riset: Deepfake, Sintesis Suara, dan Identitas Dai

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada media; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi synthetic media policy dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan provenance and consent policy.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur consent trace, detection/correction.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: impersonation.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam deepfake, sintesis suara, dan identitas dai, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah consent trace, detection/correction. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada media, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. EU AI Act, 2024; NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan provenance and consent policy. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi synthetic media policy: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara

internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada media, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator consent trace, detection/correction dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa consent trace, detection/correction; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi impersonation. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, synthetic media policy memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks media, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan

selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (EU AI Act, 2024; NIST, 2024).

Contoh penerapan riil: pada media, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas synthetic media policy, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.7 Keamanan Prompt, Data, dan Integrasi Sistem

Keamanan Prompt, Data, dan Integrasi Sistem ditempatkan sebagai unit kerja riset karena menghadapi prompt injection, data exfiltration, malicious documents, credential leakage, dan supply-chain risk. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi threat modeling + security test. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah security controls, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.8. Rencana tindakan riset: Keamanan Prompt, Data, dan Integrasi Sistem

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi security controls dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan threat modeling + security test.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur attack success rate, patch SLA.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: data leakage.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam keamanan prompt, data, dan integrasi sistem, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah attack success rate, patch SLA. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan threat modeling + security test. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi security controls: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan

indikator attack success rate, patch SLA dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa attack success rate, patch SLA; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi data leakage. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keamanan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, security controls memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus

dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (NIST, 2024).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas security controls, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

9.8 Batas Otoritas: AI Tidak Menggantikan Ulama

Batas Otoritas: AI Tidak Menggantikan Ulama ditempatkan sebagai unit kerja riset karena menentukan ruang non-fatwa, label, disclaimer, eskalasi, konsultasi manusia, dan audit untuk menjaga akuntabilitas. Fokus pada masjid/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Bunt, 2018; Whyte, 2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi policy + expert consensus. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah authority boundary policy, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 9.9. Rencana tindakan riset: Batas Otoritas: AI Tidak Menggantikan Ulama

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada masjid/lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
2. Desain	Susun spesifikasi authority boundary policy dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan policy + expert consensus.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur correct escalation, user understanding.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: automation of fatwa.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

9.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam batas otoritas: ai tidak menggantikan ulama, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah correct escalation, user understanding. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada masjid/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Bunt, 2018; Whyte, 2022 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

9.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan policy + expert consensus. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi authority boundary policy: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada masjid/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator correct escalation, user understanding dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

9.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa correct escalation, user understanding; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi automation of fatwa. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, authority boundary policy memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks masjid/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Bunt, 2018; Whyte, 2022).

Contoh penerapan riil: pada masjid/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas authority boundary policy, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 9

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan orientasi kemaslahatan dan pencegahan mudarat untuk menilai proporsionalitas penggunaan AI' sebagai contoh. Bukti dapat berasal dari ethical impact assessment, sedangkan produk atau intervensinya dapat berbentuk ethical decision memo. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti benefit-harm balance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti techno-solutionism perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguji apakah sistem memperlakukan kelompok secara adil dan tidak memperburuk ketimpangan akses atau representasi' sebagai contoh. Bukti dapat berasal dari fairness audit, sedangkan produk atau intervensinya dapat berbentuk fairness report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti error parity, access equity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti bias data dan bahasa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks Indonesia, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerapkan tujuan spesifik, minimisasi data, keamanan, retention, hak subjek, serta transparansi pemrosesan' sebagai contoh. Bukti dapat berasal dari privacy impact assessment, sedangkan produk atau intervensinya dapat berbentuk privacy notice + register. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti compliance, deletion SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti unauthorized processing perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjelaskan fungsi, batas, sumber, versi model, penggunaan AI, dan siapa manusia penanggung jawab' sebagai contoh. Bukti dapat berasal dari documentation audit, sedangkan produk atau intervensinya dapat berbentuk model/system card. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti documentation completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti false transparency perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mendeteksi keluaran tidak berdasar melalui retrieval grounding, confidence/abstention, test set, adversarial testing, dan review manusia' sebagai contoh. Bukti dapat berasal dari benchmark + red teaming, sedangkan produk atau intervensinya dapat berbentuk reliability protocol. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti groundedness, hallucination rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti silent error perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mencegah pemalsuan identitas dan penggunaan suara/wajah tokoh tanpa izin serta menetapkan pelabelan konten sintetis' sebagai contoh. Bukti dapat berasal dari provenance and consent policy, sedangkan produk atau intervensinya dapat berbentuk synthetic media policy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti consent trace, detection/correction harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti impersonation perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghadapi prompt injection, data exfiltration, malicious documents, credential leakage, dan supply-chain risk' sebagai contoh. Bukti dapat berasal dari threat modeling + security test, sedangkan produk atau intervensinya dapat berbentuk security controls. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti attack success rate, patch SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menentukan ruang non-fatwa, label, disclaimer, eskalasi, konsultasi manusia, dan audit untuk menjaga akuntabilitas' sebagai contoh. Bukti dapat berasal dari policy + expert consensus, sedangkan produk atau intervensinya dapat berbentuk authority boundary policy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti correct escalation, user understanding harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti automation of fatwa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan orientasi kemaslahatan dan pencegahan mudarat untuk menilai proporsionalitas penggunaan AI' sebagai contoh. Bukti dapat berasal dari ethical impact assessment, sedangkan produk atau intervensinya dapat berbentuk ethical decision memo. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti benefit-harm balance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti techno-solutionism perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguji apakah sistem memperlakukan kelompok secara adil dan tidak memperburuk ketimpangan akses atau representasi' sebagai contoh. Bukti dapat berasal dari fairness audit, sedangkan produk atau intervensinya dapat berbentuk fairness report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti error parity, access equity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti bias data dan bahasa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks Indonesia, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerapkan tujuan spesifik, minimisasi data, keamanan, retention, hak subjek, serta transparansi pemrosesan' sebagai contoh. Bukti dapat berasal dari privacy impact assessment, sedangkan produk atau intervensinya dapat berbentuk privacy notice + register. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti compliance, deletion SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti unauthorized processing perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjelaskan fungsi, batas, sumber, versi model, penggunaan AI, dan siapa manusia penanggung jawab' sebagai contoh. Bukti dapat berasal dari documentation audit, sedangkan produk atau intervensinya dapat berbentuk model/system card. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti documentation completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti false transparency perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mendeteksi keluaran tidak berdasar melalui retrieval grounding, confidence/abstention, test set, adversarial testing, dan review manusia' sebagai contoh. Bukti dapat berasal dari benchmark + red teaming, sedangkan produk atau intervensinya dapat berbentuk reliability protocol. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti groundedness, hallucination rate harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti silent error perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks media, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mencegah pemalsuan identitas dan penggunaan suara/wajah tokoh tanpa izin serta menetapkan pelabelan konten sintetis' sebagai contoh. Bukti dapat berasal dari provenance and consent policy, sedangkan produk atau intervensinya dapat berbentuk synthetic media policy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti consent trace, detection/correction harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti impersonation perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menghadapi prompt injection, data exfiltration, malicious documents, credential leakage, dan supply-chain risk' sebagai contoh. Bukti dapat berasal dari threat modeling + security test, sedangkan produk atau intervensinya dapat berbentuk security controls. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti attack success rate, patch SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti data leakage perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks masjid/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menentukan ruang non-fatwa, label, disclaimer, eskalasi, konsultasi manusia, dan audit untuk menjaga akuntabilitas' sebagai contoh. Bukti dapat berasal dari policy + expert consensus, sedangkan produk atau intervensinya dapat berbentuk authority boundary policy. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti correct escalation, user understanding harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti automation of fatwa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menggunakan orientasi kemaslahatan dan pencegahan mudarat untuk menilai proporsionalitas penggunaan AI' sebagai contoh. Bukti dapat berasal dari ethical impact assessment, sedangkan produk atau intervensinya dapat berbentuk ethical decision memo. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti benefit-harm balance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti techno-solutionism perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguji apakah sistem memperlakukan kelompok secara adil dan tidak memperburuk ketimpangan akses atau representasi' sebagai contoh. Bukti dapat berasal dari fairness audit, sedangkan produk atau intervensinya dapat berbentuk fairness report. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti error parity, access equity harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti bias data dan bahasa perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks Indonesia, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menerapkan tujuan spesifik, minimisasi data, keamanan, retention, hak subjek, serta transparansi pemrosesan' sebagai contoh. Bukti dapat berasal dari privacy impact assessment, sedangkan produk atau intervensinya dapat berbentuk privacy notice + register. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti compliance, deletion SLA harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti unauthorized processing perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menjelaskan fungsi, batas, sumber, versi model, penggunaan AI, dan siapa manusia penanggung jawab' sebagai contoh. Bukti dapat berasal dari documentation audit, sedangkan produk atau intervensinya dapat berbentuk model/system card. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti documentation completeness harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti false transparency perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

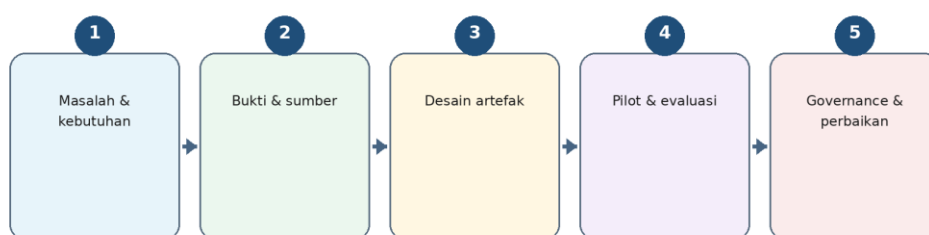
SOAL PENGEMBANGAN BAB 9

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 9 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 9, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

BAB 10 — ROADMAP R&D, EVALUASI DAMPAK, DAN MASA DEPAN DAKWAH AI

Bab 10 membangun kompetensi untuk roadmap r&d, evaluasi dampak, dan masa depan dakwah ai. Tujuan bab ini adalah membuat pembaca mampu bergerak dari konsep ke penelitian dan dari penelitian ke artefak atau intervensi yang dapat diuji. Setiap subbab menggunakan pola yang konsisten: landasan dan indikator, prosedur riset-pengembangan, kemudian evaluasi, risiko, dan keberlanjutan. Pola tersebut dimaksudkan agar buku dapat dipakai sebagai panduan kelas, manual proyek, maupun kerangka proposal penelitian.

Kerangka Operasional Bab 10



Gambar 10.1. Kerangka operasional Roadmap R&D, Evaluasi Dampak, Dan Masa Depan Dakwah Ai

Tabel 10.1. Matriks kerja utama Bab 10

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Masalah	Apa kebutuhan nyata yang belum terpenuhi?	Baseline	Problem statement	Prioritaskan/tunda
Bukti	Apa yang sudah diketahui dari literatur dan data lokal?	Literatur + data lapangan	Evidence map	Pilih pendekatan

Komponen	Pertanyaan kunci	Data	Artefak	Keputusan
Desain	Bagaimana solusi bekerja dan di mana batasnya?	Spesifikasi + korpus	Prototipe	Uji internal
Evaluasi	Apakah solusi bermanfaat, aman, dan dapat dipakai?	Metrik + wawancara + log	Laporan pilot	Go/revise/hold/stop
Governance	Siapa bertanggung jawab dan bagaimana koreksi dilakukan?	Risk register + audit trail	SOP	Scale/maintain/retire

10.1 Roadmap 90 Hari dari Masalah ke Pilot

Roadmap 90 Hari dari Masalah ke Pilot ditempatkan sebagai unit kerja riset karena menyusun sprint riset: diagnosis, literature scan, ethics, data, prototipe, benchmark, pilot, review, dan keputusan go/no-go. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Peffers et al., 2007; NIST, 2024. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi agile research + DSRM. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah 90-day pilot, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.2. Rencana tindakan riset: Roadmap 90 Hari dari Masalah ke Pilot

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi 90-day pilot dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan agile research + DSRM.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur milestone completion, risk closure.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: pilot tanpa baseline.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.1.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam roadmap 90 hari dari masalah ke pilot, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah milestone completion, risk closure. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Peffers et al., 2007; NIST, 2024 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.1.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan agile research + DSRM. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi 90-day pilot: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim

mengumpulkan indikator milestone completion, risk closure dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.1.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa milestone completion, risk closure; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi pilot tanpa baseline. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, 90-day pilot memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna

menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Peffer et al., 2007; NIST, 2024).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas 90-day pilot, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.2 Metrik Dampak: Reach, Learning, Trust, Action

Metrik Dampak: Reach, Learning, Trust, Action ditempatkan sebagai unit kerja riset karena mengembangkan metrik berlapis dari jangkauan sampai perubahan pengetahuan, kepercayaan, tindakan, dan keberlanjutan. Fokus pada semua konteks mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Glasgow et al., 1999. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi RE-AIM + logic model. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah impact scorecard, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.3. Rencana tindakan riset: Metrik Dampak: Reach, Learning, Trust, Action

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada semua konteks; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan	Masalah spesifik dan terukur

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
		sampling	
2. Desain	Susun spesifikasi impact scorecard dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan RE-AIM + logic model.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur reach, effectiveness, adoption, maintenance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: vanity metrics dan causal overclaim.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.2.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam metrik dampak: reach, learning, trust, action, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah reach, effectiveness, adoption, maintenance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada semua konteks, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi

tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Glasgow et al., 1999 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.2.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan RE-AIM + logic model. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi impact scorecard: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada semua konteks, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator reach, effectiveness, adoption, maintenance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.2.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa reach, effectiveness, adoption, maintenance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi vanity metrics dan causal overclaim. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, impact scorecard memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks semua konteks, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Glasgow et al., 1999).

Contoh penerapan riil: pada semua konteks, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas impact scorecard, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.3 TAM, UTAUT, dan Penerimaan Teknologi Dakwah

TAM, UTAUT, dan Penerimaan Teknologi Dakwah ditempatkan sebagai unit kerja riset karena mengukur usefulness, ease of use, performance expectancy, effort, social influence, dan facilitating conditions secara kontekstual. Fokus pada lembaga/sekolah mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Davis, 1989; Venkatesh et al., 2003. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi survey model testing. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah adoption model, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.4. Rencana tindakan riset: TAM, UTAUT, dan Penerimaan Teknologi Dakwah

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga/sekolah; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi adoption	Spesifikasi,	Disetujui tim

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	model dan batas fungsi.	daftar sumber, risk register	substansi dan teknis
3. Prototipe	Bangun versi minimum menggunakan survey model testing.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur construct reliability, intention/use.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: acceptance disamakan dengan kualitas.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.3.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam tam, utaut, dan penerimaan teknologi dakwah, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah construct reliability, intention/use. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga/sekolah, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar

korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Davis, 1989; Venkatesh et al., 2003 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.3.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan survey model testing. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi adoption model: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga/sekolah, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator construct reliability, intention/use dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.3.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa construct reliability, intention/use; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi acceptance disamakan dengan kualitas. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, adoption model memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga/sekolah, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Davis, 1989; Venkatesh et al., 2003).

Contoh penerapan riil: pada lembaga/sekolah, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas adoption model, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.4 CFIR dan Faktor Implementasi

CFIR dan Faktor Implementasi ditempatkan sebagai unit kerja riset karena memetakan karakteristik intervensi, setting luar/dalam, individu, dan proses implementasi untuk menjelaskan mengapa pilot berhasil atau gagal. Fokus pada lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Damschroder et al., 2022. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi CFIR-guided evaluation. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah implementation diagnosis, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.5. Rencana tindakan riset: CFIR dan Faktor Implementasi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi implementation diagnosis	Spesifikasi, daftar sumber,	Disetujui tim dan substansi

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	dan batas fungsi.	risk register	teknis
3. Prototipe	Bangun versi minimum menggunakan CFIR-guided evaluation.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur barrier/facilitator coverage.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: framework overload.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.4.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam cfir dan faktor implementasi, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah barrier/facilitator coverage. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Damschroder et al., 2022 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.4.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan CFIR-guided evaluation. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi implementation diagnosis: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator barrier/facilitator coverage dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena

R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.4.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa barrier/facilitator coverage; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi framework overload. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, implementation diagnosis memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Damschroder et al., 2022).

Contoh penerapan riil: pada lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara,

analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas implementation diagnosis, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.5 Replikasi, Skalabilitas, dan Keberlanjutan

Replikasi, Skalabilitas, dan Keberlanjutan ditempatkan sebagai unit kerja riset karena menguji apakah solusi tetap bekerja pada lokasi, kelompok, bahasa, dan periode berbeda tanpa kehilangan kualitas. Fokus pada multi-site mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain McKenney & Reeves, 2019; RE-AIM. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi replication study. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah scale-up plan, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.6. Rencana tindakan riset: Replikasi, Skalabilitas, dan Keberlanjutan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada multi-site; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi scale-up plan dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan replication study.	Version log, test set, defect log	Lulus uji internal minimum

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
4. Pilot	Uji terbatas; ukur effect stability, cost, maintenance.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: scale before evidence.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.5.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam replikasi, skalabilitas, dan keberlanjutan, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah effect stability, cost, maintenance. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada multi-site, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. McKenney & Reeves, 2019; RE-AIM memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan

kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.5.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan replication study. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi scale-up plan: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada multi-site, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator effect stability, cost, maintenance dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.5.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa effect stability, cost, maintenance; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi scale before evidence. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, scale-up plan memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks multi-site, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (McKenney & Reeves, 2019; RE-AIM).

Contoh penerapan riil: pada multi-site, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas scale-up plan, dokumentasi sumber,

instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.6 Open Science dan Reproducible Da'wah Research

Open Science dan Reproducible Da'wah Research ditempatkan sebagai unit kerja riset karena mendorong preregistration bila sesuai, berbagi instrumen, codebook, prompt version, evaluation set, dan hasil negatif tanpa membocorkan data sensitif. Fokus pada kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Page et al., 2021; NIST, 2023. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi open research workflow. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah research repository, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.7. Rencana tindakan riset: Open Science dan Reproducible Da'wah Research

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi research repository dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan open research workflow.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur reproducibility,	Dataset pilot, log insiden,	Data cukup untuk keputusan

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
	transparency.	wawancara	
5. Keputusan	Tinjau risiko: privacy conflict.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.6.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam open science dan reproducible da'wah research, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah reproducibility, transparency. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Page et al., 2021; NIST, 2023 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.6.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan open research workflow. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi research repository: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator reproducibility, transparency dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.6.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa reproducibility, transparency; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi

dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi privacy conflict. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, research repository memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Page et al., 2021; NIST, 2023).

Contoh penerapan riil: pada kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas research repository, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.7 Skenario Masa Depan: Agentic AI dan Religious Information

Skenario Masa Depan: Agentic AI dan Religious Information ditempatkan sebagai unit kerja riset karena menilai agen AI, multimodal systems, autonomous

workflows, dan implikasinya terhadap high-stakes religious information. Fokus pada lembaga/kampus mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Boettcher, 2026; Tretter et al., 2026; Stanford HAI, 2026. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi scenario planning + risk foresight. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah future scenario matrix, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.8. Rencana tindakan riset: Skenario Masa Depan: Agentic AI dan Religious Information

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada lembaga/kampus; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi future scenario matrix dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan scenario planning + risk foresight.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur preparedness, control coverage.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: delegasi berlebihan.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.7.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam skenario masa depan: agentic ai dan religious information, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah preparedness, control coverage. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi, satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada lembaga/kampus, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Boettcher, 2026; Tretter et al., 2026; Stanford HAI, 2026 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.7.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan scenario planning + risk foresight. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi future scenario matrix: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintesis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada lembaga/kampus, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator preparedness, control coverage dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.7.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa preparedness, control coverage; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi delegasi berlebihan. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, future scenario matrix memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks lembaga/kampus, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Boettcher, 2026; Tretter et al., 2026; Stanford HAI, 2026).

Contoh penerapan riil: pada lembaga/kampus, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas future scenario matrix, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

10.8 Agenda Riset 2030 dan Prinsip Penutup

Agenda Riset 2030 dan Prinsip Penutup ditempatkan sebagai unit kerja riset karena merumuskan agenda penelitian tentang efektivitas, epistemik, otoritas, bias bahasa, akses, governance, sustainability, dan hubungan manusia-AI. Fokus pada akademik/lembaga mencegah pembahasan AI menjadi abstrak: peneliti harus menyatakan siapa pengguna, kebutuhan apa yang dihadapi, proses apa yang ingin diperbaiki, serta konsekuensi jika sistem memberikan hasil yang salah. Literatur yang relevan untuk bagian ini antara lain Winter, 2026; UNESCO, 2021. Dalam buku ini, teknologi diperlakukan sebagai komponen intervensi, bukan sebagai tujuan; keberhasilan ditentukan oleh nilai layanan dan dampaknya bagi manusia.

Masalah penelitian dapat dirumuskan dengan pola: kondisi awal, kesenjangan, penyebab yang dapat dipengaruhi, intervensi yang diusulkan, outcome, serta batasan. Untuk topik ini, data yang layak dipertimbangkan meliputi Delphi + research prioritization. Data tersebut tidak otomatis harus semuanya dikumpulkan. Prinsip minimisasi mengharuskan peneliti memilih hanya data yang berhubungan dengan pertanyaan riset. Artefak yang dituju adalah agenda riset 2030, tetapi artefak baru boleh dipandang layak jika memenuhi kebutuhan pengguna, dapat dijelaskan, dapat diaudit, dan memiliki jalur koreksi ketika gagal.

Tabel 10.9. Rencana tindakan riset: Agenda Riset 2030 dan Prinsip Penutup

Tahap	Tindakan nyata	Bukti yang disimpan	Kriteria selesai
1. Diagnosis	Wawancara/observasi/survei pada akademik/lembaga; prioritaskan satu masalah.	Ringkasan baseline, consent, catatan sampling	Masalah spesifik dan terukur
2. Desain	Susun spesifikasi agenda riset 2030 dan batas fungsi.	Spesifikasi, daftar sumber, risk register	Disetujui tim dan substansi teknis
3. Prototipe	Bangun versi minimum menggunakan Delphi + research prioritization.	Version log, test set, defect log	Lulus uji internal minimum
4. Pilot	Uji terbatas; ukur priority consensus, funded studies.	Dataset pilot, log insiden, wawancara	Data cukup untuk keputusan
5. Keputusan	Tinjau risiko: agenda driven by hype.	Minutes review, keputusan dan alasan	Go/revise/hold/stop terdokumentasi

10.8.1 Landasan, konstruk, dan indikator

Konstruk utama perlu diterjemahkan menjadi definisi operasional. Dalam agenda riset 2030 dan prinsip penutup, peneliti dapat memisahkan indikator proses, keluaran langsung, outcome antara, dan outcome akhir. Indikator inti yang dapat digunakan ialah priority consensus, funded studies. Pemisahan ini penting agar angka jangkauan atau jumlah interaksi tidak secara keliru disimpulkan sebagai perubahan pemahaman dan perilaku. Satu indikator sebaiknya memiliki definisi,

satuan, sumber data, frekuensi pengukuran, penanggung jawab, dan kriteria interpretasi.

Unit analisis juga harus ditentukan sejak awal. Pada akademik/lembaga, unit dapat berupa individu pengguna, kelompok kelas, jamaah, sesi layanan, dokumen, percakapan, atau program. Pemilihan unit memengaruhi desain sampling dan analisis. Jika data percakapan dipakai, identitas sebaiknya dipisahkan dari isi analisis sejauh mungkin. Jika dokumen menjadi unit, provenance dan versi harus dipertahankan sehingga peneliti dapat menjelaskan sumber setiap keluaran yang diuji.

Hipotesis atau proposisi tidak harus selalu berbentuk hubungan statistik. Untuk penelitian pengembangan, proposisi desain dapat berbunyi bahwa transparansi sumber, batas fungsi yang jelas, dan mekanisme eskalasi manusia akan meningkatkan kepercayaan yang tepat dan menurunkan over-reliance. Proposisi tersebut kemudian dioperasionalkan ke perilaku yang dapat diamati: apakah pengguna membuka sumber, apakah sistem abstain pada pertanyaan di luar korpus, dan apakah pengguna memahami bahwa keluaran AI bukan otoritas keagamaan.

Literatur digunakan untuk membangun alasan, bukan menggantikan pengujian lokal. Winter, 2026; UNESCO, 2021 memberi kerangka awal, sementara konteks organisasi menentukan variabel tambahan seperti bahasa, tingkat literasi, kebiasaan jamaah, infrastruktur, dan kebijakan internal. Karena itu, tahap awal sebaiknya menghasilkan model konseptual satu halaman yang menghubungkan kebutuhan, mekanisme, artefak, risiko, indikator, dan keputusan yang akan diambil dari hasil evaluasi.

10.8.2 Prosedur riset dan pengembangan yang dapat diterapkan

Langkah pertama adalah diagnosis. Tim menyusun pertanyaan riset, melakukan pemetaan pemangku kepentingan, kemudian mengumpulkan baseline menggunakan Delphi + research prioritization. Hasil diagnosis ditulis sebagai daftar kebutuhan terprioritasi, bukan daftar fitur teknologi. Setiap kebutuhan diberi bukti pendukung, jumlah orang yang terdampak, tingkat urgensi, dan alternatif non-AI. Jika solusi non-AI lebih sederhana dan sama efektifnya, pilihan tersebut harus tetap dipertimbangkan.

Langkah kedua adalah desain. Tim menyusun spesifikasi agenda riset 2030: input yang boleh diterima, sumber yang dapat digunakan, fungsi yang diizinkan, fungsi yang dilarang, siapa reviewer manusia, data apa yang dicatat, dan kapan sistem harus berhenti atau mengalihkan pengguna. Spesifikasi ini berfungsi sebagai kontrak evaluasi. Dengan cara itu, perubahan model atau vendor tidak mengubah tujuan penelitian secara diam-diam.

Langkah ketiga adalah prototipe terbatas. Gunakan data non-sensitif atau data sintetis yang aman untuk pengembangan awal, lalu buat test set kecil dari kasus nyata yang sudah dianonimkan dan disetujui. Untuk setiap kasus, siapkan hasil yang diharapkan, sumber rujukan, serta kondisi kegagalan. Prototipe diuji secara internal sebelum menyentuh pengguna. Temuan dicatat sebagai defect log yang berisi jenis kesalahan, tingkat dampak, penyebab, tindakan korektif, dan status verifikasi ulang.

Langkah keempat adalah pilot lapangan. Pada akademik/lembaga, pilot sebaiknya dibatasi pada kelompok dan periode yang jelas. Pengguna diberi penjelasan mengenai tujuan, penggunaan AI, keterbatasan, serta kanal pengaduan. Tim mengumpulkan indikator priority consensus, funded studies dan data kualitatif tentang pengalaman. Evaluasi tidak hanya mencari apakah sistem 'disukai', tetapi apakah ia membantu tujuan program tanpa menimbulkan risiko yang tidak dapat diterima.

Langkah kelima adalah keputusan berbasis bukti. Kriteria go, revise, hold, atau stop ditetapkan sebelum pilot. Misalnya, sistem tidak boleh dilanjutkan bila kesalahan sumber melebihi ambang yang disetujui reviewer atau bila mekanisme eskalasi gagal pada kasus berisiko tinggi. Dokumentasi keputusan penting karena R&D yang bertanggung jawab mengakui bahwa penghentian prototipe dapat menjadi hasil penelitian yang valid.

10.8.3 Evaluasi, risiko, dan keberlanjutan

Evaluasi kuantitatif dapat memeriksa priority consensus, funded studies; evaluasi kualitatif menelusuri alasan di balik angka tersebut melalui wawancara, focus group, observasi, atau analisis keluhan. Jika terdapat log sistem, log digunakan secara proporsional dan tidak berubah menjadi pengawasan pengguna. Triangulasi dilakukan dengan membandingkan self-report, kinerja tugas, dan evidence dari sistem. Perbedaan antar-sumber data justru harus dijelaskan, bukan dihapus.

Risiko utama pada topik ini meliputi agenda driven by hype. Risk register dapat menggunakan kolom: skenario bahaya, penyebab, pihak terdampak, likelihood, severity, kontrol preventif, kontrol detektif, residual risk, pemilik risiko, dan bukti penutupan. Untuk fungsi keagamaan, kesalahan yang menyangkut sumber atau otoritas harus diberi bobot lebih tinggi daripada gangguan kosmetik pada antarmuka.

Keberlanjutan bergantung pada pemilik proses. Satu sistem yang bagus pada saat pilot dapat memburuk ketika sumber berubah, model diperbarui, pengurus berganti, atau biaya layanan meningkat. Karena itu, agenda riset 2030 memerlukan jadwal review, versi, backup, prosedur koreksi, indikator kualitas minimum, dan rencana exit. Organisasi juga perlu memastikan bahwa pengetahuan operasional tidak hanya dimiliki satu pengembang.

Pelaporan akhir harus membedakan hasil yang terbukti, temuan eksploratif, keterbatasan, dan rekomendasi. Klaim kausal hanya boleh dibuat bila desain penelitian mendukungnya. Laporan memuat konteks akademik/lembaga, karakteristik sampel, versi artefak, periode penelitian, instrumen, data yang hilang, insiden, perubahan selama pilot, dan alasan keputusan. Dengan dokumentasi ini, penelitian dapat direplikasi atau dibandingkan di lokasi lain.

Secara etis, keberhasilan berarti manfaat yang dapat dipertanggungjawabkan, bukan maksimalisasi penggunaan teknologi. Bila fungsi tertentu menuntut terlalu banyak data, memiliki risiko epistemik tinggi, atau mendorong pengguna menggantikan konsultasi manusia dengan jawaban otomatis, desain harus dipersempit. Prinsip proporsionalitas dan human oversight memberikan dasar untuk mengambil keputusan tersebut (Winter, 2026; UNESCO, 2021).

Contoh penerapan riil: pada akademik/lembaga, tim kecil beranggotakan penanggung jawab substansi agama, peneliti/evaluator, pengelola program, dan pengembang dapat menjalankan pilot selama 4-8 pekan. Minggu pertama digunakan untuk baseline dan audit sumber; minggu kedua untuk desain serta test set; minggu ketiga-keempat untuk prototipe dan uji internal; sisa periode untuk pilot, wawancara, analisis, serta review risiko. Produk minimal bukan sekadar aplikasi, tetapi paket lengkap yang terdiri atas agenda riset 2030, dokumentasi sumber, instrumen evaluasi, risk register, log versi, dan keputusan tindak lanjut. Skema ini dapat diperkecil untuk tugas mahasiswa atau diperluas untuk program organisasi.

LATIHAN DAN PEMBAHASAN BAB 10

1. Rumuskan satu masalah penelitian nyata yang sesuai dengan bab ini dan jelaskan mengapa masalah tersebut lebih penting daripada memilih teknologi terlebih dahulu.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun sprint riset: diagnosis, literature scan, ethics, data, prototipe, benchmark, pilot, review, dan keputusan go/no-go' sebagai contoh. Bukti dapat berasal dari agile research + DSRM, sedangkan produk atau intervensinya dapat berbentuk 90-day pilot. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti milestone completion, risk closure harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pilot tanpa baseline perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

2. Susun dua indikator proses dan dua indikator outcome yang dapat digunakan untuk mengevaluasi intervensi pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan metrik berlapis dari jangkauan sampai perubahan pengetahuan, kepercayaan, tindakan, dan keberlanjutan' sebagai contoh. Bukti dapat berasal dari RE-AIM + logic model, sedangkan produk atau intervensinya dapat berbentuk impact scorecard. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, effectiveness, adoption, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti vanity metrics dan causal overclaim perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

3. Jelaskan perbedaan antara data yang menarik untuk dikumpulkan dan data yang benar-benar diperlukan untuk menjawab pertanyaan riset.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengukur usefulness, ease of use, performance expectancy, effort, social influence, dan facilitating conditions secara kontekstual' sebagai contoh. Bukti dapat berasal dari survey model testing, sedangkan produk atau intervensinya dapat berbentuk adoption model. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti construct reliability, intention/use harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti acceptance disamakan dengan kualitas perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

4. Buat satu skenario kegagalan AI dan jelaskan kontrol preventif serta detektif yang diperlukan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memetakan karakteristik intervensi, setting luar/dalam, individu, dan proses implementasi untuk menjelaskan mengapa pilot berhasil atau gagal' sebagai contoh. Bukti dapat berasal dari CFIR-guided evaluation, sedangkan produk atau intervensinya dapat berbentuk implementation diagnosis. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti barrier/facilitator coverage harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti framework overload perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

5. Jelaskan kapan solusi non-AI seharusnya dipilih dibanding solusi berbasis AI.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks multi-site, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguji apakah solusi tetap bekerja pada lokasi, kelompok, bahasa, dan periode berbeda tanpa kehilangan kualitas' sebagai contoh. Bukti dapat berasal dari replication study, sedangkan produk atau intervensinya dapat berbentuk scale-up plan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti effect stability, cost, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti scale before evidence perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

6. Susun kriteria go, revise, hold, dan stop untuk sebuah pilot kecil.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mendorong preregistration bila sesuai, berbagi instrumen, codebook, prompt version, evaluation set, dan hasil negatif tanpa membocorkan data sensitif' sebagai contoh. Bukti dapat berasal dari open research workflow, sedangkan produk atau intervensinya dapat berbentuk research repository. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reproducibility, transparency harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti privacy conflict perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

7. Jelaskan bagaimana human-in-the-loop diterapkan tanpa menjadikannya sekadar formalitas review.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai agen AI, multimodal systems, autonomous workflows, dan implikasinya terhadap high-stakes religious information' sebagai contoh. Bukti dapat berasal dari scenario planning + risk foresight, sedangkan produk atau intervensinya dapat berbentuk future scenario matrix. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti preparedness, control coverage harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti delegasi berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebutkan keterbatasan dan peran reviewer manusia.

8. Rancang cara memastikan sumber keagamaan tetap dapat ditelusuri dari keluaran sistem.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks akademik/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'merumuskan agenda penelitian tentang efektivitas, epistemik, otoritas, bias bahasa, akses, governance, sustainability, dan hubungan manusia-AI' sebagai contoh. Bukti dapat berasal dari Delphi + research prioritization, sedangkan produk atau intervensinya dapat berbentuk agenda riset 2030. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti priority consensus, funded studies harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti agenda driven by hype perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

9. Jelaskan bagaimana triangulasi kuantitatif-kualitatif dapat meningkatkan kualitas kesimpulan.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun sprint riset: diagnosis, literature scan, ethics, data, prototipe, benchmark, pilot, review, dan keputusan go/no-go' sebagai contoh. Bukti dapat berasal dari agile research + DSRM, sedangkan produk atau intervensinya dapat berbentuk 90-day pilot. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti milestone completion, risk closure harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pilot tanpa baseline perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

10. Buat contoh data dictionary minimum untuk satu dataset penelitian pada bab ini.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan metrik berlapis dari jangkauan sampai perubahan pengetahuan, kepercayaan, tindakan, dan keberlanjutan' sebagai contoh. Bukti dapat berasal dari RE-AIM + logic model, sedangkan produk atau intervensinya dapat berbentuk impact scorecard. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, effectiveness, adoption, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti vanity metrics dan causal overclaim perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

11. Jelaskan risiko automation bias dan cara mengukurnya pada pengguna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengukur usefulness, ease of use, performance expectancy, effort, social influence, dan facilitating conditions secara kontekstual' sebagai contoh. Bukti dapat berasal dari survey model testing, sedangkan produk atau intervensinya dapat berbentuk adoption model. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti construct reliability, intention/use harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti acceptance disamakan dengan kualitas perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

12. Rancang satu prosedur informed consent yang singkat namun bermakna.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memetakan karakteristik intervensi, setting luar/dalam, individu, dan proses implementasi untuk menjelaskan mengapa pilot berhasil atau gagal' sebagai contoh. Bukti dapat berasal dari CFIR-guided evaluation, sedangkan produk atau intervensinya dapat berbentuk implementation diagnosis. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti barrier/facilitator coverage harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti framework overload perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

13. Jelaskan bagaimana test set dibangun agar tidak hanya berisi kasus mudah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks multi-site, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menguji apakah solusi tetap bekerja pada lokasi, kelompok, bahasa, dan periode berbeda tanpa kehilangan kualitas' sebagai contoh. Bukti dapat berasal dari replication study, sedangkan produk atau intervensinya dapat berbentuk scale-up plan. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti effect stability, cost, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti scale before evidence perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

14. Tentukan satu variabel penerimaan teknologi dan cara mengukurnya tanpa menyamakan penerimaan dengan kebenaran.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mendorong preregistration bila sesuai, berbagi instrumen, codebook, prompt version, evaluation set, dan hasil negatif tanpa membocorkan data sensitif' sebagai contoh. Bukti dapat berasal dari open research workflow, sedangkan produk atau intervensinya dapat berbentuk research repository. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reproducibility, transparency harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti privacy conflict perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

15. Jelaskan bagaimana pembaruan model atau sumber dapat mengubah validitas hasil penelitian.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/kampus, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menilai agen AI, multimodal systems, autonomous workflows, dan implikasinya terhadap high-stakes religious information' sebagai contoh. Bukti dapat berasal dari scenario planning + risk foresight, sedangkan produk atau intervensinya dapat berbentuk future scenario matrix. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti preparedness, control coverage harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti delegasi berlebihan perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

16. Buat satu contoh keputusan yang harus selalu dieskalasikan kepada manusia.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks akademik/lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'merumuskan agenda penelitian tentang efektivitas, epistemik, otoritas, bias bahasa, akses, governance, sustainability, dan hubungan manusia-AI' sebagai contoh. Bukti dapat berasal dari Delphi + research prioritization, sedangkan produk atau intervensinya dapat berbentuk agenda riset 2030. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti priority consensus, funded studies harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti agenda driven by hype perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

17. Jelaskan mengapa vanity metrics dapat menyesatkan evaluasi dakwah digital.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'menyusun sprint riset: diagnosis, literature scan, ethics, data, prototipe, benchmark, pilot, review, dan keputusan go/no-go' sebagai contoh. Bukti dapat berasal dari agile research + DSRM, sedangkan produk atau intervensinya dapat berbentuk 90-day pilot. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti milestone completion, risk closure harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti pilot tanpa baseline perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

18. Rancang mekanisme keluhan dan koreksi publik jika sistem memberi informasi yang salah.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks semua konteks, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengembangkan metrik berlapis dari jangkauan sampai perubahan pengetahuan, kepercayaan, tindakan, dan keberlanjutan' sebagai contoh. Bukti dapat berasal dari RE-AIM + logic model, sedangkan produk atau intervensinya dapat berbentuk impact scorecard. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti reach, effectiveness, adoption, maintenance harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti vanity metrics dan causal overclaim perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

19. Jelaskan komponen yang wajib ada dalam laporan pilot agar dapat direplikasi.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga/sekolah, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'mengukur usefulness, ease of use, performance expectancy, effort, social influence, dan facilitating conditions secara kontekstual' sebagai contoh. Bukti dapat berasal dari survey model testing, sedangkan produk atau intervensinya dapat berbentuk adoption model. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti construct reliability, intention/use harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti acceptance disamakan dengan kualitas perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

20. Susun satu pertanyaan penelitian lanjutan yang layak diuji pada lokasi berbeda.

Pembahasan: jawaban yang baik harus menghubungkan pertanyaan dengan konteks lembaga, tujuan yang terukur, data yang proporsional, dan kontrol risiko. Gunakan fokus 'memetakan karakteristik intervensi, setting luar/dalam, individu, dan proses implementasi untuk menjelaskan mengapa pilot berhasil atau gagal' sebagai contoh. Bukti dapat berasal dari CFIR-guided evaluation, sedangkan produk atau intervensinya dapat berbentuk implementation diagnosis. Kualitas tidak cukup dinilai dari penggunaan atau popularitas; indikator seperti barrier/facilitator coverage harus dibaca bersama temuan kualitatif dan catatan insiden. Risiko seperti framework overload perlu ditulis sebagai skenario konkret, diberi pemilik risiko, dan ditautkan ke kriteria keputusan. Jawaban juga harus menyebut keterbatasan dan peran reviewer manusia.

SOAL PENGEMBANGAN BAB 10

1. Kembangkan proposal mini 8-12 halaman yang menguji salah satu artefak pada Bab 10 menggunakan DSRM atau EDR.
2. Buat test set minimal 50 kasus untuk salah satu aplikasi Bab 10, lengkap dengan sumber rujukan dan label eskalasi.
3. Rancang pilot 6 pekan pada masjid, sekolah, kampus, atau lembaga dengan baseline, indikator, risk register, dan rencana analisis.
4. Lakukan audit etika terhadap satu alat AI yang sudah dipakai di lingkungan Anda dan petakan kesenjangannya terhadap UNESCO/NIST.
5. Susun laporan keputusan go/revise/hold/stop berdasarkan dataset simulasi yang Anda definisikan sendiri secara transparan.

PORTOFOLIO 16 PROYEK RISET DAN PENGEMBANGAN YANG SIAP DITERAPKAN

Bagian ini merangkum proyek-proyek yang dapat langsung dijadikan tugas kelompok, penelitian institusional, atau program pilot. Setiap proyek harus dimulai dengan persetujuan lembaga, penilaian risiko, dan pembatasan scope. Untuk pertanyaan hukum agama yang memerlukan otoritas, artefak wajib mengalihkan pengguna kepada manusia.

Tabel P.1. Portofolio proyek riset-pengembangan dakwah AI

Konteks	Proyek	Data utama	Metrik	Output akhir
Masjid	Chatbot layanan non-fatwa berbasis RAG	FAQ administrasi, jadwal, materi terverifikasi	Groundedness, escalation accuracy, kepuasan	Prototipe + laporan riset + SOP + instrumen + risk register
Masjid	Arsip khutbah semantik	Audio khutbah, metadata, transkrip terverifikasi	WER, search success, citation fidelity	Prototipe + laporan riset + SOP + instrumen + risk register
Masjid	Audit kebutuhan jamaah berbasis data	Survei anonim, wawancara, observasi	coverage, priority validity, program uptake	Prototipe + laporan riset + SOP + instrumen + risk register
Masjid	Literasi AI keluarga	Modul, pre-post test, contoh konten sintetis aman	learning gain, verification skill	Prototipe + laporan riset + SOP + instrumen + risk register
Sekolah	Tutor pendamping guru	Buku ajar, modul resmi, test set	learning gain, source fidelity	Prototipe + laporan riset + SOP + instrumen + risk register
Sekolah	Microlearning PAI terverifikasi	Materi ajar, soal, feedback	retention, completion,	Prototipe + laporan riset + SOP +

Konteks	Proyek	Data utama	Metrik	Output akhir
			item quality	instrumen + risk register
Sekolah	Asesmen autentik era GenAI	rubrik proses, oral defense, log revisi	mastery, fairness, workload	Prototipe + laporan riset + SOP + instrumen + risk register
Kampus	Laboratorium dakwah digital	dataset, prototipe, ethics review	rigor, utility, reproducibility	Prototipe + laporan riset + SOP + instrumen + risk register
Kampus	Digital ethnography kebutuhan mahasiswa	observasi etik, wawancara, diary	theme saturation, consent compliance	Prototipe + laporan riset + SOP + instrumen + risk register
Kampus	A/B test format pesan	dua variasi konten non-manipulatif	comprehension, voluntary engagement	Prototipe + laporan riset + SOP + instrumen + risk register
Lembaga	Pusat pengetahuan organisasi	SOP, keputusan, materi, metadata	retrieval precision, freshness	Prototipe + laporan riset + SOP + instrumen + risk register
Lembaga	Pipeline QA konten generatif	draft AI + editorial review	defect rate, review time, provenance	Prototipe + laporan riset + SOP + instrumen + risk register
Lembaga	Dashboard dampak dakwah	data program agregat	reach, effectiveness, adoption, maintenance	Prototipe + laporan riset + SOP + instrumen + risk register

Konteks	Proyek	Data utama	Metrik	Output akhir
Lembaga	Incident response AI	log kesalahan, triase, koreksi	MTTD, MTTC, recurrence	Prototipe + laporan riset + SOP + instrumen + risk register
Lembaga	Model kematangan AI	audit proses, policy, skill, data	maturity score, closure rate	Prototipe + laporan riset + SOP + instrumen + risk register
Multi-lokasi	Replikasi chatbot/knowledge hub	test set lintas bahasa/lokasi	performance stability, equity	Prototipe + laporan riset + SOP + instrumen + risk register

DAFTAR PUSTAKA

- A'yuni, Q. (2026). Transformation of Digital Da'wah in the Era of Artificial Intelligence: Analysis of Da'wah Communication Strategies on Social Media Platforms. *Wardah*, 27(1), 86–98. <https://doi.org/10.19109/wardah.v27i1.35047>
- Whyte, S. A. (2022). Islamic Religious Authority in Cyberspace: A Qualitative Study of Muslim Religious Actors in Australia. *Religions*, 13(1), 69. <https://doi.org/10.3390/rel13010069>
- Al-Qur'an al-Karim.
- Alhammad, N., Awae, F., & Yussuf, A. (2025). Integrating Artificial Intelligence in Islamic Education: A Review on Pedagogical Approaches and Learning Outcomes. *International Journal of Academic Research in Business and Social Sciences*, 15(7). <https://doi.org/10.6007/IJARBS/v15-i7/25947>
- Arif, M., Nor Abd Aziz, M. K., Hassan, H., Fikri, Y., Zekkari, M., & Rismawati, B. V. (2026). Artificial Intelligence in Islamic Religious Education: A PRISMA-Based Review and the SEI-CHAT Framework. *Online Learning in Educational Research*, 6(1), 65–85. <https://doi.org/10.58524/oler.v6i1.993>
- Association of Internet Researchers. (2019). *Internet Research: Ethical Guidelines 3.0*.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of FAccT '21*, 610-623. <https://doi.org/10.1145/3442188.3445922>
- Boettcher, S. J. (2026). *Artificial Intelligence and Spirituality: Relating Religious Information and Religious Knowledge*. Springer VS. <https://doi.org/10.1007/978-3-031-93058-4>
- Braun, V., & Clarke, V. (2021). *Thematic Analysis: A Practical Guide*. SAGE.
- Bucher, T. (2018). *If...Then: Algorithmic Power and Politics*. Oxford University Press.
- Bunt, G. R. (2003). *Islam in the Digital Age: E-Jihad, Online Fatwas and Cyber Islamic Environments*. Pluto Press.
- Bunt, G. R. (2018). *Hashtag Islam: How Cyber-Islamic Environments Are Transforming Religious Authority*. University of North Carolina Press. <https://doi.org/10.5149/northcarolina/9781469643168.001.0001>
- Campbell, H. A., & Bellar, W. (2023). *Digital Religion: The Basics*. Routledge.
- Campbell, H. A., & Tsuria, R. (Eds.). (2022). *Digital Religion: Understanding Religious Practice in Digital Media* (2nd ed.). Routledge.
- Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107, 1753-1820.
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). SAGE.

- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE.
- Damschroder, L. J., Reardon, C. M., Widerquist, M. A. O., & Lowery, J. (2022). The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science*, 17, 75. <https://doi.org/10.1186/s13012-022-01245-0>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, 32(1), 5-8. <https://doi.org/10.3102/0013189X032001005>
- DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). SAGE.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (4th ed.). Wiley.
- European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.
- Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health*, 89(9), 1322-1327. <https://doi.org/10.2105/AJPH.89.9.1322>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105.
- Indonesia, Kementerian Komunikasi dan Informatika. (2023). Surat Edaran Menteri Komunikasi dan Informatika Nomor 9 Tahun 2023 tentang Etika Kecerdasan Artifisial.
- Indonesia. (2022). Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi.
- ISO/IEC. (2023). ISO/IEC 42001:2023 Information technology - Artificial intelligence - Management system. International Organization for Standardization.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248. <https://doi.org/10.1145/3571730>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>

- Kemmis, S. (2009). Action research as a practice-based practice. *Educational Action Research*, 17(3), 463-474. <https://doi.org/10.1080/09650790903093284>
- Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The Action Research Planner: Doing Critical Participatory Action Research*. Springer. <https://doi.org/10.1007/978-981-4560-67-2>
- Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology* (4th ed.). SAGE.
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., et al. (2018). The science of fake news. *Science*, 359(6380), 1094-1096. <https://doi.org/10.1126/science.aao2998>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems*, 33.
- McKenney, S., & Reeves, T. C. (2019). *Conducting Educational Design Research* (2nd ed.). Routledge.
- Miao, F., & Holmes, W. (2023). *Guidance for Generative AI in Education and Research*. UNESCO.
- Muhammad, K. (2025). Artificial Intelligence in Islamic Religious Education: A Systematic Review of Applications, Challenges, and Future Directions. *Social Science Academic*, 3(2), 255–264. <https://doi.org/10.37680/ssa.9990>
- Muslim ibn al-Hajjaj. *Sahih Muslim*.
- NIST. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. National Institute of Standards and Technology.
- NIST. (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1)*. <https://doi.org/10.6028/NIST.AI.600-1>
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.
- O'Neil, C. (2016). *Weapons of Math Destruction*. Crown.
- OECD. (2024). *OECD AI Principles* (updated 2024). Organisation for Economic Co-operation and Development.
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., et al. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of*

- Management Information Systems, 24(3), 45-77. <https://doi.org/10.2753/MIS0742-1222240302>
- Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). Free Press.
- Stanford Institute for Human-Centered Artificial Intelligence. (2026). *The 2026 AI Index Report*. Stanford University.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Tretter, M., Scholz, M., Puzio, A., & Simmerlein, J. (Eds.). (2026). *Artificial Intelligence in Religious Practice: Bots, Robots, and Generative AI*. Springer Nature. <https://doi.org/10.1007/978-3-032-23721-7>
- UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. UNESCO.
- UNESCO. (2024). *AI Competency Framework for Teachers*. UNESCO.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. Council of Europe.
- Winter, D. (2026). Religious ethics in the age of AI: A comparative study of faith-based approaches to AI governance. *AI and Ethics*, 6, Article 298. <https://doi.org/10.1007/s43681-026-01156-6>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39. <https://doi.org/10.1186/s41239-019-0171-0>

GLOSARIUM

A

Abstention: Kemampuan sistem untuk tidak menjawab ketika bukti, confidence, atau ruang lingkup tidak memadai.

AI literacy: Kompetensi memahami, menggunakan, mengevaluasi, dan mengawasi AI secara kritis dan bertanggung jawab.

Audit trail: Jejak terdokumentasi mengenai sumber, perubahan, keputusan, dan tindakan dalam siklus hidup sistem.

B

Baseline: Pengukuran kondisi awal sebelum intervensi.

Benchmark: Kumpulan tugas dan kriteria yang dipakai untuk menguji kinerja secara konsisten.

Bias algoritmik: Pola kesalahan sistematis yang menghasilkan perlakuan atau keluaran tidak adil.

C

CFIR: Consolidated Framework for Implementation Research, kerangka untuk menganalisis faktor implementasi.

Chunk: Potongan dokumen yang diindeks dan diambil kembali dalam sistem retrieval.

Consent: Persetujuan yang diberikan setelah informasi yang relevan dipahami oleh partisipan.

D

Data dictionary: Dokumen definisi variabel, tipe data, nilai, sumber, dan aturan transformasi.

Data minimization: Prinsip mengumpulkan hanya data yang diperlukan untuk tujuan yang jelas.

Design Science Research: Pendekatan penelitian yang menghasilkan dan mengevaluasi artefak untuk menyelesaikan masalah nyata.

E

Embedding: Representasi numerik yang memungkinkan perbandingan kemiripan semantik.

Escalation: Pengalihan kasus dari sistem otomatis kepada manusia yang berwenang.

Explainability: Kemampuan memberikan informasi yang memadai mengenai alasan atau proses sistem.

F

Fairness: Prinsip memperlakukan kelompok secara adil dan memeriksa ketimpangan dampak.

Fidelity: Derajat kesesuaian implementasi dengan desain atau sumber yang ditetapkan.

Focus group: Diskusi kelompok terarah untuk menggali pengalaman, persepsi, dan alasan.

G

Generative AI: AI yang menghasilkan konten baru seperti teks, audio, gambar, atau kode berdasarkan pola yang dipelajari.

Gold set: Kumpulan kasus uji dengan jawaban atau label rujukan yang telah diverifikasi ahli.

Groundedness: Derajat keterikatan keluaran model pada bukti atau dokumen yang disediakan.

H

Hallucination: Keluaran model yang terdengar meyakinkan tetapi tidak didukung sumber atau fakta yang benar.

Human-in-the-loop: Desain yang menempatkan manusia pada titik review atau keputusan tertentu.

Human oversight: Pengawasan manusia untuk memastikan tanggung jawab akhir tidak diserahkan kepada AI.

I

Impact assessment: Penilaian sistematis atas manfaat, risiko, dan dampak kepada pihak yang terdampak.

Incident log: Catatan kesalahan atau kejadian yang berpotensi menimbulkan kerugian.

Inter-rater agreement: Tingkat kesepakatan antarpemilai pada proses coding atau scoring.

J

Jangkauan (reach): Proporsi atau jumlah kelompok sasaran yang benar-benar terpapar atau menerima layanan.

Jejak sumber: Informasi yang memungkinkan pengguna menelusuri asal materi hingga dokumen referensi.

K

Knowledge base: Kumpulan pengetahuan terstruktur yang menjadi sumber layanan pencarian atau tanya jawab.

Korpus: Kumpulan dokumen yang dikurasi untuk penelitian atau sistem bahasa.

Kriteria stop: Ambang yang menyebabkan pilot atau sistem harus dihentikan.

L

Large Language Model: Model bahasa berskala besar yang memprediksi dan menghasilkan urutan bahasa.

Logic model: Representasi hubungan antara input, aktivitas, output, outcome, dan dampak.

Log sistem: Catatan kejadian teknis atau interaksi yang dipakai untuk monitoring dan evaluasi.

M

Mad'u: Pihak atau kelompok sasaran dakwah yang harus dipahami kebutuhan dan konteksnya.

Metadata: Data yang menjelaskan data lain, seperti sumber, tanggal, versi, tema, dan bahasa.

Mixed methods: Desain penelitian yang mengintegrasikan pendekatan kuantitatif dan kualitatif.

N

Natural Language Processing: Bidang AI untuk memproses, memahami, dan menghasilkan bahasa manusia.

Needs assessment: Proses sistematis untuk mengidentifikasi kebutuhan, kesenjangan, dan prioritas.

Non-fatwa assistant: Asisten informasi yang secara eksplisit tidak mengambil peran pemberi fatwa.

O

Ontologi: Model konseptual formal tentang entitas dan hubungan di suatu domain.

Outcome: Perubahan yang terjadi setelah intervensi, berbeda dari output yang hanya mencatat aktivitas.

Over-reliance: Ketergantungan berlebihan pada keluaran sistem tanpa verifikasi yang memadai.

P

Participatory Action Research: Penelitian partisipatif yang melibatkan anggota komunitas dalam siklus tindakan dan refleksi.

Prompt: Instruksi dan konteks yang diberikan kepada model untuk mengarahkan tugas.

Provenance: Informasi asal-usul data, dokumen, atau keluaran yang memungkinkan keterlacakan.

Q

Quality assurance: Rangkaian kontrol untuk memastikan keluaran memenuhi standar sebelum dipublikasikan.

Qualitative coding: Proses memberi kode pada data kualitatif untuk menemukan pola dan tema.

R

RAG: Retrieval-Augmented Generation; pendekatan yang menggabungkan retrieval dokumen dan generasi.

RE-AIM: Kerangka evaluasi Reach, Effectiveness, Adoption, Implementation, dan Maintenance.

Risk register: Daftar terstruktur skenario risiko, kontrol, pemilik, dan residual risk.

S

Semantic search: Pencarian berdasarkan kemiripan makna, bukan hanya kecocokan kata.

Sosioteknis: Cara pandang bahwa teknologi, manusia, organisasi, dan norma membentuk satu sistem.

System card: Dokumentasi fungsi, batasan, evaluasi, risiko, dan penggunaan sistem.

T

Tabayyun: Prinsip verifikasi informasi sebelum menerima atau menyebarkannya.

TAM: Technology Acceptance Model.

Theory of Change: Penjelasan mekanisme bagaimana aktivitas diharapkan menghasilkan outcome.

U

Usability: Derajat kemudahan, efisiensi, dan kepuasan penggunaan sistem dalam konteks tertentu.

UTAUT: Unified Theory of Acceptance and Use of Technology.

V

Validation: Proses menguji bahwa instrumen, data, atau artefak sesuai tujuan dan kriteria.

Versioning: Pengelolaan versi data, model, prompt, dan dokumen agar perubahan dapat dilacak.

W

Workflow: Urutan kerja yang mendefinisikan siapa melakukan apa, kapan, dan dengan bukti apa.

Word Error Rate: Metrik kesalahan transkripsi otomatis pada speech recognition.

X

XAI: Explainable Artificial Intelligence; pendekatan untuk meningkatkan keterjelasan perilaku atau keluaran AI.

Y

Yield evaluasi: Proporsi kasus uji yang menghasilkan keluaran layak setelah semua kontrol kualitas diterapkan.

Z

Zero-shot: Kemampuan model melakukan tugas tanpa contoh spesifik dalam prompt; tetap memerlukan evaluasi sebelum penggunaan nyata.

LAMPIRAN INSTRUMEN RISET DAN PENGEMBANGAN

A. Lembar Analisis Kebutuhan

Instrumen A. Lembar Analisis Kebutuhan dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran A. Template Lembar Analisis Kebutuhan

Komponen	Isian/pertanyaan	Catatan/bukti
Masalah atau kebutuhan utama		
Kelompok terdampak		
Bukti baseline		
Penyebab yang diduga		
Alternatif non-AI		
Manfaat potensial AI		
Risiko utama		
Data minimum		

Komponen	Isian/pertanyaan	Catatan/bukti
Pemilik proses		
Keputusan prioritas		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

B. Pedoman Wawancara Pengguna

Instrumen B. Pedoman Wawancara Pengguna dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran B. Template Pedoman Wawancara Pengguna

Komponen	Isian/pertanyaan	Catatan/bukti
Pengalaman saat ini		
Kesulitan terbesar		
Sumber informasi yang dipercaya		
Cara memverifikasi informasi		
Pengalaman menggunakan AI		

Komponen	Isian/pertanyaan	Catatan/bukti
Harapan terhadap layanan		
Kekhawatiran privasi		
Kapan ingin berbicara dengan manusia		
Bentuk penjelasan sumber yang disukai		
Saran perbaikan		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

C. Kuesioner Penerimaan Teknologi

Instrumen C. Kuesioner Penerimaan Teknologi dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran C. Template Kuesioner Penerimaan Teknologi

Komponen	Isian/pertanyaan	Catatan/bukti
Sistem membantu tugas saya		
Sistem mudah dipelajari		

Komponen	Isian/pertanyaan	Catatan/bukti
Saya memahami batas fungsi sistem		
Saya dapat menemukan sumber jawaban		
Saya tahu kapan harus meminta bantuan manusia		
Saya merasa kontrol privasi memadai		
Saya percaya sistem hanya untuk fungsi yang dijelaskan		
Saya akan menggunakan sistem untuk tugas yang sesuai		
Saya dapat melaporkan kesalahan		
Saya tidak menganggap sistem sebagai otoritas agama		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

D. Rubrik Fidelity Sumber

Instrumen D. Rubrik Fidelity Sumber dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran D. Template Rubrik Fidelity Sumber

Komponen	Isian/pertanyaan	Catatan/bukti
Kutipan sesuai sumber		
Konteks tidak berubah		
Nomor/rujukan benar		
Terjemahan tidak menambah makna		
Sumber ditampilkan		
Perbedaan pendapat tidak disamakan		
Batas scope terlihat		
Kondisi abstain berjalan		
Reviewer tercatat		
Versi dokumen tercatat		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika

data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

E. Test Set RAG

Instrumen E. Test Set RAG dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran E. Template Test Set RAG

Komponen	Isian/pertanyaan	Catatan/bukti
ID kasus		
Pertanyaan		
Kategori		
Jawaban rujukan		
Dokumen sumber		
Passage relevan		
Harus abstain?		
Harus eskalasi?		
Kriteria skor		

Komponen	Isian/pertanyaan	Catatan/bukti
		
Catatan reviewer		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

F. Checklist Hallucination

Instrumen F. Checklist Hallucination dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran F. Template Checklist Hallucination

Komponen	Isian/pertanyaan	Catatan/bukti
Apakah ada klaim tanpa sumber?		
Apakah sumber benar-benar mendukung klaim?		
Apakah nama kitab/nomor hadis akurat?		
Apakah model mengarang konsensus?		
Apakah model menyembunyikan ketidakpastian?		
Apakah kutipan terlalu		

Komponen	Isian/pertanyaan	Catatan/bukti
panjang/tidak perlu?		
Apakah konteks pertanyaan diabaikan?		
Apakah ada informasi yang sudah kedaluwarsa?		
Apakah jawaban melewati batas non-fatwa?		
Apakah perlu koreksi publik?		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

G. Risk Register

Instrumen G. Risk Register dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran G. Template Risk Register

Komponen	Isian/pertanyaan	Catatan/bukti
ID risiko		
Skenario bahaya		
Penyebab		

Komponen	Isian/pertanyaan	Catatan/bukti
		
Pihak terdampak		
Likelihood		
Severity		
Kontrol preventif		
Kontrol detektif		
Residual risk		
Pemilik/target penutupan		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

H. Data Dictionary

Instrumen H. Data Dictionary dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran H. Template Data Dictionary

Komponen	Isian/pertanyaan	Catatan/bukti
Nama variabel		
Definisi		
Tipe data		
Nilai/kategori		
Sumber		
Tujuan		
Sensitivitas		
Retention		
Transformasi		
Catatan kualitas		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika

data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

I. Model/System Card Ringkas

Instrumen I. Model/System Card Ringkas dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran I. Template Model/System Card Ringkas

Komponen	Isian/pertanyaan	Catatan/bukti
Nama sistem/versi		
Tujuan		
Pengguna		
Fungsi diizinkan		
Fungsi dilarang		
Data/korpus		
Metrik evaluasi		
Known limitations		

Komponen	Isian/pertanyaan	Catatan/bukti
Human oversight		
Kontak incident		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

J. Laporan Pilot

Instrumen J. Laporan Pilot dapat disalin dan diadaptasi. Setiap perubahan sebaiknya diberi nomor versi, tanggal, penanggung jawab, dan alasan perubahan. Jangan mengumpulkan data identitas jika tidak diperlukan untuk tujuan penelitian.

Tabel Lampiran J. Template Laporan Pilot

Komponen	Isian/pertanyaan	Catatan/bukti
Ringkasan eksekutif		
Masalah dan tujuan		
Konteks dan partisipan		
Desain intervensi		
Metode dan instrumen		
Hasil kuantitatif		

Komponen	Isian/pertanyaan	Catatan/bukti
		
Hasil kualitatif		
Insiden dan risiko		
Keterbatasan		
Keputusan dan rencana tindak lanjut		

Petunjuk analisis: setelah instrumen terisi, ringkas temuan ke dalam evidence table. Bedakan temuan faktual, interpretasi peneliti, dan keputusan desain. Jika data saling bertentangan, jangan memilih hanya data yang mendukung prototipe; jadikan perbedaan itu sebagai bahan revisi hipotesis atau rancangan.

- KASMUI -
Dosen Kimia, Komputasi, IT, AI UNNES

<https://hisabmu.com/>
<https://kasmui.cloud/>

Ketua PCM Gunungpati 2	Anggota Majelis Tabligh PDM Kota Semarang & PWM Jawa Tengah	Tim Pengembang Software KHGT MTT PP Muhammadiyah	Praktisi Ilmu Falak
---------------------------	---	--	------------------------