

REFINED

Research Synthesis Framework (RSF)

28 prompts for querying a research knowledge base (RKB) built from a selected corpus of papers

VERDICT: Strong backbone, not yet perfect.

The original covers the research lifecycle well, from inventory to reference list. Its main weaknesses are the absence of grounding rules, weak citation anchoring, sequencing errors, and several topic-specific steps presented as universal. The refinement keeps the structure and corrects these points.

Roadmap: six phases



Tag legend

CORE	Run for every topic.	MODULE	Use if relevant to the topic.	PROPOSAL	Reasons beyond the corpus.	NEW	Added in this version.
-------------	----------------------	---------------	-------------------------------	-----------------	----------------------------	------------	------------------------

1. Diagnosis: what was improved and why

Eleven issues were identified in the original framework. Each is paired with its refinement.

#	Issue	Problem in the original	Refinement
1	Tool-specific, no grounding rules	Prompts name one product and contain no guardrails. Retrieval-based tools answer from a limited set of retrieved passages (Lewis et al., 2020) and can fabricate references (Walters & Wilder, 2023).	Added a tool-agnostic Grounding Protocol: sources only, “Not reported in the corpus”, [Inferred] labels.
2	No persistent citation anchors	Author and year are re-derived at every step and can be misread.	Step 1 builds a Source Register (IDs and APA keys) that all later steps reuse.
3	Prompts are not self-contained	Step 4 refers to “the previous step”; this breaks if the session is reset or steps are skipped.	Each prompt names the step and output it depends on.
4	Sequence problems	The matrix (19) and empirical data (17) came after the gaps (14). The abstract (25) preceded the review (26). Methodological critique (22) followed the objectives (16).	Reordered into six phases that follow the logic of the research process.
5	Overclaiming language	“What the collective literature proves” overstates what a small corpus can show.	Cautious verbs and a four-level evidence-strength rating.
6	Vague or unbounded outputs	“Top dominant themes” and “every single document” give no limits; long lists can exceed response limits.	Set theme counts, columns, word limits, and batching. Added checkable listings before counts.
7	Topic-specific steps treated as universal	Use cases, technical specifications, barriers, and risks assume an implementation-type topic.	Steps are tagged CORE or MODULE; modules apply only where relevant.
8	Ungrounded generative steps not flagged	Significance steps ask for reasoning that goes beyond the corpus.	Phase E is tagged PROPOSAL; corpus evidence is separated from [Inferred] argument.

#	Issue	Problem in the original	Refinement
9	Gap analysis lacks traceability	Gaps rely only on authors' statements; RQs and objectives do not link to them.	Gap IDs with type, basis, and priority; each RQ and objective maps back to a Gap ID.
10	No quality assurance	Nothing checks coverage, citation matching, or reference accuracy.	New Step 27 audit; Step 28 uses the register and flags [Verify].
11	Whole review in one prompt	A single prompt tends to yield source-by-source summaries.	Review drafted theme by theme, then linked to the gaps.

2. Grounding Protocol: paste this once at the start

Replace **[TOPIC]** with the research topic. Paste the protocol again whenever the session is reset. It applies to all 28 steps.

GROUNDING PROTOCOL

You are a research analyst working exclusively with the sources uploaded to this knowledge base on **[TOPIC]**. Follow these rules in every answer.

1. Use only the uploaded sources. Never use outside knowledge.
2. Cite every claim in APA 7th in-text format and add the Source ID, e.g., (Sharma & Lee, 2021) [S07].
3. If the corpus does not address a point, write: *Not reported in the corpus*.
4. Never invent quotations, statistics, page numbers, or DOIs.
5. Label your own interpretations **[Inferred]** and keep them separate from what authors report.
6. State how many sources support each claim.
7. Use formal, objective, gender-neutral academic language.

Why each rule matters

Rules 1 and 3	prevent outside knowledge from being blended into corpus-based claims.
Rule 2	keeps every claim traceable to the Source Register built in Step 1.
Rule 4	guards against fabricated citations, which are a documented risk (Walters & Wilder, 2023).
Rules 5 and 6	let the reader separate reported evidence from interpretation, and weigh how many sources stand behind a claim.

3. Step map: original to refined

All 27 original steps are retained. One step was added (27), and the total is now 28.

New	Refined step	Tag	Original and change
1	Corpus Inventory and Source Register	CORE	Original 1 (revised)
2	Methodological Taxonomy	CORE	Original 2 (revised)
3	Temporal and Contextual Profile	CORE	Original 8 (expanded)
4	Evidence Synthesis Matrix	CORE	Original 19 (moved earlier, revised)
5	Inductive Thematic Coding	CORE	Original 3 (revised)
6	Descriptive Evidence Extraction	CORE	Original 4 (revised)
7	Empirical Data Synthesis	CORE	Original 17 (moved earlier, revised)
8	Evidence Statements with Strength Rating	CORE	Original 5 (revised)
9	Theoretical Framework Extraction	CORE	Original 6 (revised)
10	Conceptual Framework Synthesis	CORE	Original 7 (revised)
11	Comparative Cross-Source Analysis	CORE	Original 13 (revised)
12	Use Cases and Implementation Models	MODULE	Original 9 (revised)
13	Technical and Structural Analysis	MODULE	Original 10 (revised)
14	Barriers and Enablers	MODULE	Original 11 (expanded)
15	Legal, Ethical, and Governance Constraints	MODULE	Original 12 (revised)
16	Risk and Mitigation Mapping	MODULE	Original 18 (revised)
17	Cross-Sector and Interdisciplinary Transfer	MODULE	Original 24 (revised)
18	Methodological Critique and Innovation	CORE	Original 22 (moved earlier, revised)
19	Research Gaps Identification	CORE	Original 14 (revised)
20	Research Questions (RQs)	CORE	Original 15 (revised)
21	Research Objectives (OBs)	CORE	Original 16 (revised)
22	Theoretical Significance	PROPOSAL	Original 20 (revised)
23	Practical and Managerial Significance	PROPOSAL	Original 21 (revised)
24	Policy and Societal Impact	PROPOSAL	Original 23 (revised)
25	Literature Review Synthesis (Theme by Theme)	CORE	Original 26 (split by theme)
26	Summary Abstract Draft	CORE	Original 25 (moved after review)
27	Coverage and Citation Audit	NEW	New step
28	Final Reference List Generation	CORE	Original 27 (revised)

4. The refined framework: 28 steps

Run the Grounding Protocol first. Then paste each prompt in order. Where a prompt contains [X], substitute the theme name.

PHASE A | MAP THE CORPUS

Steps 1–4. Build a verified inventory and profile of the corpus.

STEP 1 Corpus Inventory and Source Register

CORE

PROMPT TO PASTE

Scan all uploaded sources and produce a Source Register as a table with these columns: Source ID (S01, S02, ...), APA 7th in-text key (Author, Year), Full Title, Author(s), Year, Publication Venue, DOI (write “Not available” if absent), Document Type (journal article, conference paper, report, thesis, book chapter, other), Country or Context Studied, and Core Research Objective (one sentence). List every document. If the output is too long, present sources in batches of 15 and continue when I reply “next”. Finish with the total number of sources and ask me to confirm it.

Purpose: Create a complete, verifiable inventory. The Source IDs and citation keys anchor every later step.

Expected output: Source Register table and confirmed total count.

Origin: Original 1 (revised)

STEP 2 Methodological Taxonomy

CORE

PROMPT TO PASTE

Using the Source Register, classify every source by research design: quantitative, qualitative, mixed-methods, systematic or scoping review, bibliometric or scientometric, conceptual or theoretical, technical design or system development, case study, or other (specify). First list each Source ID with its design, data source, and sample or dataset size. Then report counts and percentages per category in a table. Conclude with the dominant designs and the methodological blind spots (designs that are absent or rare).

Purpose: Reveal predominant designs and blind spots. Listing each source first makes the counts checkable.

Expected output: Classification table, distribution table, short synopsis.

Origin: Original 2 (revised)

STEP 3 Temporal and Contextual Profile

CORE

PROMPT TO PASTE

Profile the corpus by (a) publication year, (b) country or region, (c) sector or population studied, and (d) key terminology used. Present counts in tables. Then describe, in chronological phases, how the research focus shifted from the earliest to the most recent sources, citing Source IDs for each phase. Note any under-represented years, regions, or contexts.

Purpose: Track conceptual change over time and expose coverage imbalances that later feed the gap analysis.

Expected output: Frequency tables and phase-wise narrative.

Origin: Original 8 (expanded)

STEP 4 Evidence Synthesis Matrix

CORE

PROMPT TO PASTE

Generate a Markdown table covering every source, one row per source, with these columns: Source ID, Author(s) and Year (APA 7th key), Core Focus, Methodology, Sample or Data, Key Findings, and Stated Limitations. Keep each cell under 30 words and write "Not reported" for missing information.

Purpose: Provide a compact scaffold early in the workflow, so that later steps can be cross-checked against it.

Expected output: Literature matrix ready for manuscripts.

Origin: Original 19 (moved earlier, revised)

PHASE B | DISCOVER AND EXTRACT

Steps 5–9. Identify themes, then extract and grade the evidence.

STEP 5 Inductive Thematic Coding

CORE

PROMPT TO PASTE

Perform an inductive thematic analysis of all sources without imposing predefined categories. Identify 5 to 8 dominant themes. For each theme provide: a label, a one-sentence definition, 2 to 4 sub-themes, the number of supporting sources, and the Source IDs. Then list any minority themes found in fewer than three sources. Present the result as a table (the Theme Codebook).

Purpose: Discover patterns from the text itself (Braun & Clarke, 2006) and fix a numbered codebook for reuse.

Expected output: Theme Codebook table.

Origin: Original 3 (revised)

STEP 6 Descriptive Evidence Extraction

CORE

PROMPT TO PASTE

Using the Theme Codebook from Step 5, extract for each theme the main findings, key metrics, and core arguments reported by the authors. Present each theme as a table with the columns: Finding or Argument, Metric (if any), Source (APA 7th in-text citation), and Design and Context. Use only information stated in the sources.

Purpose: Gather raw evidence that is traceable to individual sources.

Expected output: One evidence table per theme.

Origin: Original 4 (revised)

STEP 7 Empirical Data Synthesis

CORE

PROMPT TO PASTE

Extract and tabulate all quantitative results reported across the sources (for example, survey results, performance benchmarks, effect sizes, sample sizes, and percentages), and summarise qualitative case outcomes. Use the columns: Metric or Outcome, Value, Context or Sample, and Source (APA 7th). Do not calculate or estimate figures that the sources do not report. Flag measures that are not comparable.

Purpose: Strengthen the evidentiary weight of the review with hard metrics, before gaps are drawn.

Expected output: Metrics table with comparability flags.

Origin: Original 17 (moved earlier, revised)

STEP 8 Evidence Statements with Strength Rating

CORE

PROMPT TO PASTE

For each theme, write 2 to 3 concise declarative Evidence Statements that summarise what the literature reports, supported by APA 7th in-text citations. Use cautious verbs such as “indicates” or “suggests”, not “proves”. After each statement, rate the strength of evidence as Strong (many sources, consistent findings, robust designs), Moderate, Limited (fewer than three sources), or Conflicting, and justify the rating in one line.

Purpose: Convert evidence into precise assertions and show how much weight each can bear (cf. Guyatt et al., 2008).

Expected output: Evidence Statements with rating and rationale.

Origin: Original 5 (revised)

PROMPT TO PASTE

Identify all theories, models, and conceptual lenses used in the sources. Present a table with: Theory or Model, Sources Applying It (APA 7th), How It Is Applied, and Explicit or Implicit. State how many sources use no explicit theory. Conclude by noting which theories dominate and which are under-used.

Purpose: Position the study within established theory and expose atheoretical work.

Expected output: Theory table and short synopsis.

Origin: Original 6 (revised)

PHASE C | MODEL AND COMPARE

Steps 10–17. Synthesise frameworks, debates, and topic-specific modules.

STEP 10 Conceptual Framework Synthesis

CORE

PROMPT TO PASTE

Based on the variables, relationships, and mechanisms reported in the sources, synthesise a conceptual framework. First, tabulate independent variables, mediating or moderating factors, and dependent outcomes, each with supporting citations. Then list each proposed relationship as “A → B” with the sources that support it, and mark it [Author-stated] or [Inferred]. Finally, give a text description that I can convert into a diagram.

Purpose: Map cause-and-effect relations without letting the tool invent unsupported links.

Expected output: Variable table, relationship list, diagram description.

Origin: Original 7 (revised)

STEP 11 Comparative Cross-Source Analysis

CORE

PROMPT TO PASTE

Identify where sources contradict one another, propose competing solutions, or reach different conclusions. For each debate provide: the issue, Position A with citations, Position B with citations, and possible explanations for the difference (for example, context, design, or period), marked [Inferred].

Purpose: Expose scholarly debates and separate genuine disagreement from differences in context or design.

Expected output: Debate table.

Origin: Original 13 (revised)

STEP 12 Use Cases and Implementation Models

MODULE

PROMPT TO PASTE

Extract all distinct applications, implementation models, and real-world use cases described in the sources. Present a table: Use Case, Context or Sector, Implementation Model, Reported Outcome, and Source (APA 7th). Distinguish implemented cases from proposed or theoretical models.

Purpose: Pinpoint concrete deployments versus conceptual proposals.

Expected output: Use-case table.

Origin: Original 9 (revised)

STEP 13 Technical and Structural Analysis

MODULE

PROMPT TO PASTE

Examine the technical or structural specifications, architectures, or frameworks discussed in the sources. Compare how authors evaluate alternatives in a table: Option, Strengths, Weaknesses, Evidence Reported, and Source (APA 7th).

Purpose: Evaluate the structural or technical options proposed for implementation.

Expected output: Comparison table of options.

Origin: Original 10 (revised)

STEP 14 Barriers and Enablers

MODULE

PROMPT TO PASTE

Synthesise the institutional, financial, technical, cultural, and administrative barriers to adoption or implementation, and the enabling factors, reported in the sources. Group them by category, cite in APA 7th, and state how many sources mention each.

Purpose: Identify friction points and success factors; the original covered barriers only.

Expected output: Categorised barrier and enabler tables.

Origin: Original 11 (expanded)

STEP 15**Legal, Ethical, and Governance Constraints****MODULE****PROMPT TO PASTE**

Analyse what the sources state about legal frameworks, compliance, data privacy, ethics, and governance constraints. Cite in APA 7th. For any topic the corpus does not address, write “Not reported in the corpus.”

Purpose: Highlight regulatory and ethical constraints without importing outside law.

Expected output: Structured summary with explicit gaps.

Origin: Original 12 (revised)

STEP 16**Risk and Mitigation Mapping****MODULE****PROMPT TO PASTE**

Compile the risks identified in the literature and the mitigation strategies proposed by the authors in a table: Risk, Category, Proposed Mitigation, and Source (APA 7th). Flag risks for which no mitigation is proposed.

Purpose: Provide a risk-management overview for decision-makers.

Expected output: Risk register with unmitigated risks flagged.

Origin: Original 18 (revised)

STEP 17**Cross-Sector and Interdisciplinary Transfer****MODULE****PROMPT TO PASTE**

Identify where sources borrow lessons, frameworks, or practices from other sectors or disciplines. Present a table: Origin Sector or Discipline, Concept or Practice Transferred, How It Is Adapted, Reported Result, and Source (APA 7th). Then note adjacent fields that the corpus does not draw upon, marked [Inferred].

Purpose: Show knowledge transfer and suggest cross-disciplinary links.

Expected output: Transfer table and short note on missing fields.

Origin: Original 24 (revised)

PHASE D | GAPS, QUESTIONS, OBJECTIVES

Steps 18–21. Move from the evidence to a justified research design.

STEP 18 Methodological Critique and Innovation

CORE

PROMPT TO PASTE

Evaluate the limitations of the research designs used across the sources (for example, sampling, data quality, validity, and generalisability). Summarise the recurring weaknesses with citations. Then propose how a new study could use a stronger or mixed-methods design. Keep corpus-based critique separate from recommendations marked [Inferred].

Purpose: Critique past designs before gaps and objectives are set, so that design choices are justified.

Expected output: Critique summary and design recommendations.

Origin: Original 22 (moved earlier, revised)

STEP 19 Research Gaps Identification

CORE

PROMPT TO PASTE

Review the limitations, conclusions, and future-work statements in all sources, together with the coverage patterns found in Steps 2, 3, and 5. Present a Gap Table with the columns: Gap ID (G1, G2, ...), Gap Description, Gap Type (knowledge, methodological, theoretical, contextual, population, or empirical), Basis ([Author-stated] or [Inferred from coverage]), Supporting Sources (APA 7th), and Priority (High, Medium, or Low, with a one-line rationale).

Purpose: Justify novelty with traceable gaps, drawn from author statements and from corpus coverage.

Expected output: Gap Table with IDs.

Origin: Original 14 (revised)

STEP 20 Research Questions (RQs)

CORE

PROMPT TO PASTE

Formulate 3 to 5 rigorous, researchable Research Questions grounded in the Gap Table. For each provide: RQ number, the question, the Gap ID(s) it addresses, the question type (descriptive, explanatory, comparative, or evaluative), and a one-sentence justification citing relevant sources. Add a testable hypothesis only where the design is quantitative.

Purpose: Ensure each RQ traces back to a documented gap.

Expected output: RQ table linked to Gap IDs.

Origin: Original 15 (revised)

STEP 21 Research Objectives (OBs)

CORE

PROMPT TO PASTE

Translate each Research Question into one primary objective and 2 to 3 secondary objectives, each beginning with an action verb (for example, examine, compare, develop). Present a table that maps RQ to objectives and to the methodological approach suggested in Step 18.

Purpose: Create operational goals that guide the methodology.

Expected output: RQ-to-objective mapping table.

Origin: Original 16 (revised)

PHASE E | SIGNIFICANCE AND IMPACT

Steps 22–24. Proposal-oriented steps that reason beyond the corpus.

STEP 22 Theoretical Significance

PROPOSAL

PROMPT TO PASTE

Explain how addressing the gaps in Step 19 would advance theory or extend the models identified in Steps 9 and 10. Support each contribution with citations to the corpus, and mark any argument that goes beyond the corpus as [Inferred].

Purpose: Articulate scholarly contribution while separating evidence from argument.

Expected output: Contribution statements with evidence labels.

Origin: Original 20 (revised)

STEP 23 Practical and Managerial Significance

PROPOSAL

PROMPT TO PASTE

Synthesise the actionable value of the corpus findings for practitioners and decision-makers. Present a table: Stakeholder, Actionable Insight, Supporting Evidence (APA 7th), and Implementation Consideration. Mark extrapolations as [Inferred].

Purpose: Establish real-world utility with a clear evidence trail.

Expected output: Stakeholder insight table.

Origin: Original 21 (revised)

STEP 24 Policy and Societal Impact

PROPOSAL

PROMPT TO PASTE

Examine what the sources report about wider policy and societal implications, such as equity, public trust, accessibility, and governance. Cite in APA 7th. Where the corpus is silent, write “Not reported in the corpus.”

Purpose: Broaden the scope to macro-level implications without unsupported speculation.

Expected output: Structured summary with explicit silences.

Origin: Original 23 (revised)

PHASE F | WRITE-UP AND QUALITY ASSURANCE

Steps 25–28. Draft, audit, and reference the synthesis.

STEP 25

Literature Review Synthesis (Theme by Theme)

CORE

PROMPT TO PASTE

Act as an expert academic writer. Using the Theme Codebook (Step 5), the Evidence Statements (Step 8), and the Source Register (Step 1), draft the literature review section for Theme [X] only, in 400 to 600 words. Use formal academic prose and critical synthesis (compare and contrast sources rather than summarising them one by one), with APA 7th in-text citations. End with one sentence linking to the next theme. Repeat for each theme. Then request a brief introduction and a concluding synthesis that links all themes to the research gaps.

Purpose: A theme-by-theme approach produces deeper, concept-centred synthesis (Webster & Watson, 2002).

Expected output: Draft review sections, introduction, and conclusion.

Origin: Original 26 (split by theme)

STEP 26 Summary Abstract Draft

CORE

PROMPT TO PASTE

Draft a structured abstract of 200 to 250 words (Background, Purpose, Method, Key Findings, Implications) that summarises the literature review produced in Step 25. Use only points supported by earlier outputs and keep in-text citations to a minimum.

Purpose: Summarise the completed review; the abstract now follows the review.

Expected output: Structured abstract.

Origin: Original 25 (moved after review)

STEP 27 Coverage and Citation Audit

NEW

PROMPT TO PASTE

Audit the synthesis produced in this session. (a) List any sources in the Source Register that were never cited. (b) List every in-text citation that does not match a Source Register entry. (c) Identify claims that lack a citation, or that rest on a single source but are presented as consensus. (d) Flag any statement that may not be supported by the corpus.

Purpose: Detect omissions and misattributions before submission. Manual spot-checks remain necessary.

Expected output: Audit report with corrective actions.

Origin: New step

STEP 28 Final Reference List Generation

CORE

PROMPT TO PASTE

Compile an alphabetised reference list of only those sources cited in the synthesis, using the metadata in the Source Register. Format strictly according to APA 7th (hanging indent, sentence-case article titles, italicised journal names and volumes, DOI as <https://doi.org/...>). Mark any entry with missing metadata as [Verify]. Do not add or infer DOIs, volumes, or page numbers that are not in the sources.

Purpose: Produce a reference list that matches the in-text citations. Validate it against Scopus or Crossref.

Expected output: APA 7th reference list with [Verify] flags.

Origin: Original 27 (revised)

5. Recommendations for use

1	Paste the Grounding Protocol first, with the topic filled in, and repeat it after any session reset.
2	Run all CORE steps in order. Add MODULE steps only where they suit the topic.
3	Verify counts from Steps 1 to 3 manually against the corpus, because retrieval-based tools can miss documents.
4	Save each output outside the tool, especially the Source Register, Theme Codebook, and Gap Table.
5	For large corpora, run steps by Source ID range (for example, S01-S25) and merge the results.
6	Spot-check a sample of citations against the original papers, at least 10% as a working minimum.
7	Validate every reference and DOI in Scopus or Crossref before submission.
8	Treat Phase E outputs as draft argument, not as evidence.

References

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924–926. <https://doi.org/10.1136/bmj.39489.470347.AD>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.
- Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, 13, Article 14045. <https://doi.org/10.1038/s41598-023-41032-5>
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), xiii–xxiii.