

PROTOCOL

Guide AI-Supported Narrative Review of Papers/ State-of-the-Art

Anabela Mesquita, IPP (Porto, Portugal)

(ENGLISH VERSION)

July 2026

INDEX

1. Introduction	¡Error! Marcador no definido.
Quick Start (for fast execution)	¡Error! Marcador no definido.
2. WORKING MODELS	¡Error! Marcador no definido.
Working Models	¡Error! Marcador no definido.
3. TOOLS (keep it tight).....	¡Error! Marcador no definido.
Step 1 — Define the Research Question	¡Error! Marcador no definido.
Step 2 — Paper Selection (Human-led)	¡Error! Marcador no definido.
Step 3 — Structured Evidence Extraction (Critical Step).....	¡Error! Marcador no definido.
Step 4 — Build the Synthesis Matrix.....	¡Error! Marcador no definido.
Step 5 — Initial Thematic Synthesis.....	¡Error! Marcador no definido.
Step 6 — Validate and Refine the Themes (New Step)	¡Error! Marcador no definido.
Step 7 — Prepare the Paper Outline	¡Error! Marcador no definido.
Step 8 — Draft the Paper	¡Error! Marcador no definido.
Model A (Recommended)	¡Error! Marcador no definido.
Model B (Co-writing)	¡Error! Marcador no definido.
Step 9 — Develop the Critical Discussion	¡Error! Marcador no definido.
Step 10 — Evidence Audit and Quality Control	¡Error! Marcador no definido.
4. FINAL STRUCTURE (expected output)	¡Error! Marcador no definido.
5. COMMON FAILURE POINTS (be explicit with colleagues).	¡Error! Marcador no definido.
6. OPTIONAL ENHANCEMENT	¡Error! Marcador no definido.
7. BOTTOM LINE	¡Error! Marcador no definido.

1. Introduction

This protocol provides a structured approach for preparing a short narrative review or state-of-the-art paper (up to 3000 words) based on a corpus of 10 to 20 academic articles. Its purpose is to guide researchers in using artificial intelligence in a controlled and rigorous way, ensuring that the final output reflects genuine synthesis and critical analysis rather than a collection of disconnected summaries.

The protocol is built around a clear workflow that covers all stages of the process: defining a focused research question, selecting relevant papers, extracting structured information, organising evidence in a synthesis matrix, identifying themes, and developing a coherent and critical narrative. It distinguishes between two modes of AI use. In Model A, AI acts as an assistant supporting analysis and structuring, while the researcher retains full control over interpretation. In Model B, AI can support drafting, but always under strict validation by the researcher.

The main objective is to improve both the efficiency and the quality of literature-based work. On the one hand, it reduces time spent on repetitive tasks such as information extraction and

initial structuring. On the other hand, it reinforces essential academic practices, particularly transparency, traceability of claims, and critical engagement with the literature.

This protocol is designed for researchers who need to produce concise, well-structured papers accessible to both academic audiences and informed readers outside the field, without compromising analytical depth or methodological rigour.

Since recent generations of AI models are capable of performing increasingly sophisticated reasoning tasks, researchers may be tempted to delegate interpretation and synthesis to AI. This protocol explicitly discourages that approach. AI should support analytical work, but the intellectual responsibility for identifying patterns, weighing evidence, resolving contradictions and constructing arguments remains entirely with the researcher.

Quick Start (for fast execution)

This section provides a simplified sequence for rapid execution. Each step corresponds to a detailed section in the protocol.

- 1 Define research question
- 2 Select papers
- 3 Structured extractions
- 4 Build synthesis matrix
- 5 Initial thematic analysis
- 6 Validate themes against the papers
- 7 Refine synthesis
- 8 Draft sections
- 9 Critical discussions
- 10 Evidence audit and quality control

2. WORKING MODELS

Working Models

This protocol supports two modes of using AI, which differ in level of autonomy and risk. In Model A, AI is used as an assistant to support structured extraction, organisation of information, and identification of themes, while the researcher remains fully responsible for interpretation and synthesis. This is the preferred approach for ensuring rigor and control. In Model B, AI takes a more active role as a co-writer, assisting in drafting sections of the paper. However, this model requires strict validation, as the researcher must verify all claims, ensure alignment with the source material, and strengthen the critical dimension. The choice between models depends on the researcher's experience, time constraints, and risk tolerance, but in all cases, responsibility for the final output remains entirely human.

Level	AI task	Human responsibility
Level 1	Information extraction	Verify accuracy
Level 2	Comparison across studies	Validate patterns
Level 3	Interpretation and synthesis	Make final analytical judgement

Model A — AI as Assistant (recommended baseline)

AI is used for:

- Structured extraction
- Thematic suggestions
- Draft support

Humans are responsible for:

- Interpretation
- Synthesis
- Critical analysis

Model B — AI as Co-writer (controlled use)

AI is used for:

- Drafting sections
- Rewriting and improving flow

Human must:

- Validate all content against sources
- Rewrite critical sections
- Remove generic or fabricated statements

3. TOOLS (keep it tight)

The protocol relies on a small set of tools selected for their functional value rather than variety. ChatGPT is used to support structured extraction, synthesis, and drafting, always based on researcher-provided inputs. NotebookLM can be used to work across multiple documents simultaneously, improving traceability and reducing the risk of unsupported claims. Microsoft Excel is essential for building the synthesis matrix, which underpins the entire analytical process. Finally, Zotero ensures proper organisation of references and accurate citations. The choice of tools is secondary to the structure of the process, but misuse or over-reliance on any single tool will compromise the quality of the final output.

Use only what adds value:

- ChatGPT → extraction, synthesis, drafting
- NotebookLM → multi-document grounding (strong for Model A)
- Microsoft Excel → synthesis matrix (non-negotiable)
- Zotero → references and citation control
- ChatGPT Projects (good for iterative synthesis)

- Claude (strong long-context synthesis)
- Gemini (very large context windows)

Optional:

- Do NOT ask AI to identify consensus without verifying supporting studies.
- Do NOT accept phrases such as "the literature suggests" unless the supporting evidence has been checked.
- Do NOT allow AI to introduce concepts that are not grounded in the reviewed papers.
- Do NOT confuse fluent writing with scientific synthesis.

4. WORKFLOW

The workflow follows a structured but iterative process designed to ensure that the final paper is analytical, coherent, and firmly grounded in evidence rather than a collection of disconnected summaries. Although the stages are presented in sequence, researchers should expect to move back and forth between them as their understanding of the literature evolves. Revisiting earlier steps is not a sign of poor planning but an essential part of producing a rigorous literature review.

The process begins by defining a clear and focused research question. This question provides the analytical framework for the entire review and determines which studies should be included, which evidence is relevant, and which themes are worth exploring. If the research question is too broad or poorly defined, the resulting review will lack focus and analytical depth.

Once the research question has been established, the researcher selects between 10 and 20 relevant papers. Selection should be based on their direct contribution to answering the research question rather than on convenience or availability. Preference should be given to papers that provide conceptual, empirical or methodological value. Including marginally relevant studies usually weakens the overall synthesis by introducing unnecessary variation.

Each paper is then analysed individually using a structured extraction template. Rather than producing free-form summaries, the researcher extracts comparable information from every paper, including research objectives, methodology, context, principal findings, limitations and key concepts. Standardising the extraction process facilitates systematic comparison across studies and reduces the risk of overlooking important evidence.

The extracted information is organised into a synthesis matrix, typically using a spreadsheet. The matrix constitutes the analytical foundation of the review because it makes similarities, differences, recurring concepts and methodological patterns visible across the literature. Skipping this step almost invariably results in descriptive writing instead of genuine synthesis.

Based on the synthesis matrix, the researcher performs an initial thematic analysis by identifying recurring topics, patterns, areas of agreement, contradictions and gaps in the

literature. AI can assist by suggesting possible thematic groupings, but these suggestions should always be treated as hypotheses rather than conclusions. Every proposed theme must be verified against the original papers to ensure that it accurately reflects the available evidence.

This validation stage is essential. As themes emerge, the researcher should return to the original papers and the synthesis matrix to confirm whether the evidence genuinely supports the proposed interpretation. It is common for themes to be refined, merged, divided or even discarded during this iterative process. The objective is not simply to identify recurring topics but to construct themes that accurately represent the body of evidence.

Only after the thematic structure has been validated should the researcher prepare the outline of the paper. At this stage, the synthesis has already been developed. The outline simply organises the existing analytical structure into a coherent manuscript, typically including an introduction, a brief description of the review process, thematic sections, a critical discussion, identification of research gaps and a conclusion.

The drafting stage follows. AI may assist in writing individual sections, improving clarity or restructuring paragraphs, but the text should always be generated from the validated synthesis rather than directly from the individual papers. The role of AI is to support communication, not interpretation. The researcher remains fully responsible for ensuring that every statement accurately reflects the underlying evidence and that the manuscript integrates studies instead of merely describing them.

A critical discussion is then developed. This is the defining feature of a high-quality state-of-the-art review. Rather than summarising the literature, the researcher evaluates it by identifying where evidence converges, where it diverges, which areas are over-researched, which questions remain insufficiently explored, and which methodological limitations recur across studies. Particular attention should also be given to the strength of the available evidence, avoiding broad conclusions when only limited support exists.

Before the manuscript is considered complete, an evidence audit should be conducted. Every major claim should be checked against the reviewed literature to verify that it is explicitly supported by the analysed papers. Statements such as "the literature suggests", "research shows" or "evidence indicates" should only be used when the supporting evidence has been systematically verified. This final verification substantially reduces unsupported generalisations and improves the credibility of the review.

Finally, the paper undergoes a comprehensive quality control process. This includes verifying logical consistency, confirming that all claims are evidence-based, checking references and citations, improving clarity and readability, and ensuring that no unsupported content introduced during AI-assisted drafting remains in the final version. The objective is to ensure

that AI has enhanced efficiency without compromising the scientific rigour, transparency or critical quality of the review.

Step 1 — Define the Research Question

A well-defined research question is the foundation of the entire review. It determines which papers should be included, which evidence is relevant, and how the literature will be analysed. A question that is too broad will produce a descriptive review rather than a focused synthesis.

Prompt

I am preparing a narrative review based on 10–20 academic papers.

Help me refine the following topic into a precise and researchable question suitable for a state-of-the-art review.

Topic: [INSERT TOPIC]

Requirements

- Produce three alternative research questions.
- Each question should be specific, researchable and suitable for thematic synthesis.
- Avoid broad or descriptive formulations.
- Briefly explain the advantages and limitations of each option.
- Recommend the strongest research question and justify your choice.

Step 2 — Paper Selection (Human-led)

Paper selection remains a human responsibility. AI may assist in identifying potentially relevant studies, but the final decision should always be made by the researcher.

Minimum selection criteria

- Direct relevance to the research question.
- Conceptual, theoretical or empirical contribution.
- Methodological diversity where appropriate.
- Avoid unnecessary redundancy.
- Prefer recent publications (last 5–10 years unless justified).
- Avoid multiple papers based on the same dataset unless they contribute different perspectives.
- Ensure explicit alignment with the review objective.

Before proceeding, confirm that the selected corpus provides sufficient breadth and depth to answer the research question.

Step 3 — Structured Evidence Extraction (Critical Step)

Each paper must be analysed using the same extraction template. Standardisation allows meaningful comparison across studies and forms the basis for subsequent synthesis.

Prompt

Read the following academic paper and extract structured information.

[UPLOAD OR PASTE PAPER]

Provide the following information.

1. Full reference (APA if available)
2. Research objective
3. Research question(s) (if explicitly stated)
4. Methodology (design, sample, data collection and analysis)
5. Principal findings (bullet points only, no interpretation)
6. Main concepts or variables
7. Limitations identified by the authors
8. Keywords and potential themes
9. Contradictions, tensions or inconsistencies identified within the paper
10. Strength of the evidence (High / Moderate / Limited). Justify your assessment in one sentence, considering the study design, methodology, sample, and robustness of the findings.

Rules

- Do not infer information that is not explicitly supported by the paper.
- If information is unavailable, write "Not specified".
- Keep the extraction factual and concise.
- Separate factual reporting from interpretation.

Step 4 — Build the Synthesis Matrix

The synthesis matrix is the analytical core of the review. It transforms isolated paper summaries into comparable evidence.

Minimum columns should include:

- * Author (Year)
- Research objective
- Methodology
- Context / Sample
- Key findings
- Main concepts

- Emerging themes
- Limitations
- Evidence type
- Strength of evidence

Additional columns may be added depending on the review.

Evidence type helps distinguish the nature of each contribution (e.g., empirical study, conceptual paper, systematic review, narrative review, case study or policy paper). Recording this information makes it easier to interpret differences across studies and to assess the overall balance of the evidence.

Step 5 — Initial Thematic Synthesis

The objective is to identify patterns across studies rather than summarise them individually. AI may suggest possible themes, but these should always be treated as preliminary hypotheses that require validation against the original papers. At this stage, the emphasis is on identifying meaningful patterns while avoiding unsupported generalisations.

Prompt

Based exclusively on the following synthesis matrix:

[INSERT MATRIX]

Perform the following tasks:

1. Identify four to six potential themes emerging from the reviewed studies.
2. Group the papers under each theme.
3. For each theme, identify:
 - areas of convergence;
 - areas of divergence;
 - unresolved questions;
 - gaps in the literature.
4. Assess the overall strength of evidence supporting each theme (Strong / Moderate / Limited) and briefly justify your assessment.
5. Indicate whether the evidence supporting each theme comes primarily from:
 - empirical studies;
 - conceptual papers;
 - review papers;
 - case studies;
 - policy or position papers;
 - a mixed body of evidence.
6. Identify any themes that appear plausible but are not sufficiently supported by the available evidence. Explain why these themes should be treated with caution.

Rules

- Base every conclusion exclusively on the supplied evidence.
- Do not extrapolate beyond the reviewed studies.
- Clearly distinguish established findings from tentative observations.
- Distinguish evidence-supported themes from AI-generated hypotheses.
- Avoid presenting limited evidence as scientific consensus.

Step 6 — Validate and Refine the Themes (New Step)

Before writing begins, every proposed theme should be validated against the original studies.

The objective is to confirm that the identified themes genuinely emerge from the literature rather than from AI-generated generalisations.

For each theme, verify:

- Which papers explicitly support it?
- Which papers contradict it? If contradictory findings exist, can they be explained by differences in methodology, context, sample or theoretical perspective?
- Which papers provide only partial support?
- Is the available evidence sufficient to justify the theme?
- Should any themes be merged, divided or discarded?

Only validated themes should be retained for writing.

Step 7 — Prepare the Paper Outline

Once the thematic structure has been validated, prepare a logical outline for the paper.

Prompt

Based on the validated thematic synthesis, prepare an outline for a narrative review (maximum 3000 words).

Include:

1. Introduction
2. Method
3. Thematic analysis
4. Critical discussion
5. Research gaps
6. Conclusion

Ensure:

- Logical progression.
- Balanced section lengths.
- Clear analytical structure.
- Appropriate transitions between sections.

Step 8 — Draft the Paper

Model A (Recommended)

Write the following section using only the validated synthesis provided.

[INSERT VALIDATED SYNTHESIS]

Requirements:

- No invented claims or references.
- Formal academic style.
- Integrate studies rather than describe them individually.
- Explicitly indicate disagreements where they exist.
- Avoid generic statements such as "the literature suggests" unless the evidence clearly supports them.
- Compare studies whenever appropriate.
- Explain why findings converge or diverge.
- Reflect the relative strength of the available evidence. Length: **[Specify]**

Model B (Co-writing)

Draft the following section using the validated synthesis below.

[INSERT VALIDATED MATERIAL]

Requirements:

- Maintain academic rigour.
- Do not introduce unsupported information.
- Clearly distinguish consensus from isolated findings.
- Suggest two or three ways to strengthen the critical analysis.

All AI-generated text must subsequently undergo manual revision.

Step 9 — Develop the Critical Discussion

This section distinguishes a genuine state-of-the-art review from a descriptive literature review.

Prompt

Develop a critical discussion based exclusively on the validated synthesis.

Address the following questions.

1. Which topics dominate the literature?
2. Which topics remain underexplored?
3. Where do studies disagree?
4. Which methodological limitations recur?
5. How robust is the available evidence?
6. Which assumptions are widely accepted but insufficiently supported?

Rules:

- Be analytical rather than descriptive.
- Avoid speculation.
- Support every conclusion with evidence from the reviewed studies.

Step 10 — Evidence Audit and Quality Control

Before finalising the manuscript, systematically verify every major claim.

Prompt

Review the following manuscript.

Check for:

1. Unsupported claims.
2. Logical inconsistencies.
3. Overgeneralisations.
4. Missing connections between studies.
5. Weak transitions.
6. Claims based on insufficient evidence.
7. Statements introduced by AI that are not traceable to the reviewed papers.

For every problematic sentence:

- Quote the sentence.
- Explain the problem.
- Suggest a corrected version.
- Indicate which paper(s) support the revision.

Finally, verify:

- Are all major claims explicitly supported?
- Are contradictory findings acknowledged?
- Is the strength of evidence appropriately represented?
- Has any unsupported information been introduced during drafting?
- Do any conclusions go beyond what the available evidence warrants? If so, identify them and suggest a more evidence-based formulation?

Only after successfully completing this audit should the manuscript be considered ready for submission.



5. FINAL STRUCTURE (expected output)

This section defines the expected organisation of the paper, ensuring consistency, clarity, and logical progression of ideas. By following a standard structure, the researcher avoids producing a fragmented text and instead develops a coherent argument from introduction to conclusion. Each section has a specific role: the introduction frames the problem, the method ensures transparency, the thematic analysis presents the core findings, and the discussion and gaps sections provide critical depth. This structure also supports readability for both academic and general audiences, making the paper easier to follow and evaluate. Deviating from it typically results in unclear argumentation and weak integration of sources.

1. Introduction
2. Method
3. Thematic analysis (3–5 sections)
4. Discussion
5. Gaps and future research
6. Conclusion

Evaluation Criteria

- Clarity of research question
- Quality of synthesis (not summaries)
- Strength of thematic structure
- Depth of critical discussion
- Use of evidence
- Coherence and readability

6. COMMON FAILURE POINTS (be explicit with colleagues)

This section highlights the most frequent mistakes that compromise the quality of the final paper. Its purpose is preventive: to make researchers aware of where the process typically breaks down so they can avoid repeating those errors. Common issues include relying on AI-generated summaries without proper synthesis, skipping the synthesis matrix, working with an unclear research question, and failing to develop a critical discussion. Another frequent problem is accepting AI outputs at face value without verification. Being aware of these pitfalls helps maintain rigor and ensures that the final work reflects genuine analysis rather than superficial compilation.

- Treating AI summaries as synthesis
- Skipping the matrix
- Writing without a clear research question
- Accepting AI-generated text without verification
- Producing descriptive instead of critical work
- Creating themes before comparing studies.
- Treating AI-generated patterns as evidence.
- Giving equal weight to all studies regardless of methodological quality.

- Confusing fluent writing with scientific reasoning.

7. OPTIONAL ENHANCEMENT

This section presents additional practices that can strengthen the analysis without being mandatory for completing the protocol. Tools such as NotebookLM can be used to work across multiple documents simultaneously, allowing researchers to query the full set of papers and validate emerging themes more efficiently. These tools can help detect connections, confirm patterns, and reduce the risk of overlooking relevant evidence. However, they do not replace the need for a structured synthesis matrix or critical judgment. Their role is to support verification and depth, not to automate interpretation.

If they use NotebookLM:

- Upload all papers
- Ask cross-document questions
- Use it to validate themes

8. BOTTOM LINE

If you follow this:

- You produce a credible short paper

If you skip:

- extraction structure
- synthesis matrix
- critical discussion

You will produce something that looks academic but isn't.