AI Researchers
Embedding Responsible AI in Research

AI READINESS TOOLKIT

UC1-01 Use case 1: Literature review and synthesis

From literature mapping to hypothesis development with generative AI

Using AI to surface research streams and connect them to a testable hypothesis

AT A GLANCE

A researcher uses generative AI to map the literature around a topic, identify the most relevant research streams and link their findings to testable hypotheses. The case shows how to prompt for research streams, check every suggested source against the original and decide which evidence supports each hypothesis. AI speeds up the mapping. The researcher selects the sources, judges the evidence and owns the interpretation.

TYPE
Case study

FORMAT
PDF

CONTRIBUTING PARTNER
UPF Barcelona School of Management

WHO IT IS FOR
Researchers and doctoral candidates building a literature review

FRAMEWORK BARRIERS
Trust in AI, Know-how, This task

HUMAN-ONLY BOUNDARY

Every source and interpretation is checked against the originals. The final analysis stays with the researcher.

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Context

A literature review is much more than a collection of summaries of previous studies. Researchers need to identify the main research streams within a body of literature, understand how these streams relate to one another, recognise areas of agreement and disagreement, identify theoretical and empirical gaps, and use this synthesis to justify their research questions and hypotheses.

This is particularly challenging for early-stage researchers. After collecting and reading a substantial number of papers, they may have extensive notes and summaries but still struggle to develop a coherent conceptual structure. Literature reviews can consequently become descriptive—"Author A found X, while Author B found Y"—rather than analytical and theory-driven.

Generative AI can support this stage of the research process. When provided with a controlled corpus of academic literature, GenAI tools can help researchers identify recurring themes, compare theoretical perspectives, cluster findings into research streams, detect contradictory or underexplored findings, and explore how the available evidence supports—or challenges—proposed hypotheses.

However, GenAI also creates important methodological risks. It may oversimplify heterogeneous research streams, create artificial relationships between studies, overlook methodological differences, misrepresent findings, or generate apparently plausible claims that are not supported by the source literature.

This case study therefore focuses on the critical and responsible use of GenAI as a literature-synthesis assistant rather than as an autonomous literature reviewer.

Case Scenario

Participants take the role of a researcher working on a research project for which an initial literature search has already been completed.

They receive a small but sufficiently diverse corpus of academic papers, or structured extracts from these papers, together with:

an initial research question;

a preliminary literature review;

a set of proposed hypotheses;

a bibliography and/or evidence table containing the selected studies.

The initial literature review contains a realistic problem: although it covers relevant studies, it is primarily descriptive and does not clearly reveal the intellectual structure of the literature. Moreover, the relationship between the reviewed evidence and the proposed hypotheses is weak or implicit.

Participants are asked to use GenAI to refine rather than replace the literature review.

Their challenge is to transform the existing material into a structured synthesis that answers four questions:

What are the main research streams? → What does each stream tell us? → Where are the gaps, tensions, or inconsistencies? → How does this evidence justify the proposed hypotheses?

The exercise therefore reproduces a common stage of an actual research project: moving from literature collection and description to theoretical synthesis and hypothesis development.

Main Learning Goals

The case study will enable participants to learn how to use GenAI critically and effectively to:

> Identify and characterise research streams within an existing corpus of academic literature.
>
> Compare and connect research streams, recognising complementarities, contradictions, boundary conditions, and research gaps.
>
> Assess the evidential basis of hypotheses, determining which findings support, contradict, qualify, or do not address each proposed relationship.
>
> Strengthen the theoretical argument behind hypotheses, using GenAI to explore possible causal mechanisms and logical connections while ensuring that claims remain grounded in the literature.
>
> Use iterative prompting, progressively refining GenAI outputs rather than relying on a single prompt.
>
> Recognise responsible-use issues, including hallucinated evidence, invented citations, overgeneralisation, confirmation bias, loss of nuance, and inappropriate delegation of scholarly judgement.

Case Study Stages

STAGE 1 – DIAGNOSE THE INITIAL LITERATURE REVIEW

Participants first read the titles and abstracts of the supplied literature review without using GenAI.

They identify weaknesses such as excessive description, unclear thematic organisation, missing connections between studies, unsupported generalisations, or weak links between the literature and hypotheses.

This establishes a baseline against which the value of GenAI can later be assessed.

STAGE 2 – IDENTIFY THE MAIN RESEARCH STREAMS

Participants provide the controlled literature corpus or evidence table to GenAI and ask it to propose possible ways of organising the literature.

Rather than simply requesting “the main themes,” participants experiment with different analytical criteria, for example:

> theoretical perspectives;
>
> dependent and independent variables;
>
> methodological approaches;
>
> convergent versus contradictory findings.

Participants compare alternative classifications and determine which structure provides the strongest basis for their research question.

Output: A validated map of approximately 3–5 research streams, including the defining characteristics and key studies associated with each stream.

STAGE 3 – SYNTHESISE FINDINGS WITHIN AND ACROSS STREAMS

For each research stream, participants use GenAI to identify:

> central concepts;

> dominant theoretical arguments;
>
> main empirical findings;
>
> areas of consensus;
>
> contradictory findings;

They then ask GenAI to compare streams rather than analyse them independently.

Output: A literature synthesis matrix showing how the research streams relate to one another.

STAGE 4 – IDENTIFY GAPS AND TENSIONS

Participants ask GenAI to identify potential gaps based exclusively on the supplied evidence.

A critical component of the exercise is distinguishing between:

> demonstrated gaps, which can reasonably be inferred from the corpus;
>
> possible gaps, which require further literature searching;
>
> AI-generated speculation, which cannot be substantiated.

This teaches researchers that an AI statement such as “few studies have examined X” cannot be accepted as evidence of a literature gap without verification.

STAGE 5 – LINK LITERATURE FINDINGS TO THE PROPOSED HYPOTHESES

Participants create an evidence-to-hypothesis matrix.

For each hypothesis, GenAI helps identify studies that:

> directly support the hypothesis;
>
> indirectly support its underlying mechanism;
>
> provide contradictory evidence;
>
> are relevant but insufficient to establish the proposed relationship.

Participants then evaluate whether each hypothesis is adequately grounded in the literature.

Output: A traceable Hypothesis–Evidence Matrix.

Content Structure

RESEARCH CONTEXT AND RESEARCH QUESTION

RESEARCH CONTEXT

Generative AI (GenAI) tools are rapidly becoming part of academic research practices. Researchers can use tools such as ChatGPT and other AI-based assistants to support multiple stages of the research process, including searching and exploring literature, summarising academic papers, generating and refining research questions, analysing qualitative or quantitative information, supporting coding tasks, improving academic writing, and communicating research findings.

Despite the potential benefits, researchers differ considerably in their willingness to adopt GenAI. Some perceive these tools as useful mechanisms for increasing productivity, supporting creativity, or performing research tasks more efficiently. Others are reluctant to use them because of concerns about reliability, hallucinations, privacy,

research integrity, transparency, copyright, or institutional rules.

Understanding what drives researchers' adoption of GenAI is therefore relevant for universities and research organisations seeking to develop policies, training programs, and support mechanisms for the responsible use of AI in research.

A research team would like to investigate the factors influencing researchers' intention to use GenAI as part of their research activities.

MAIN RESEARCH QUESTION

RQ: What factors influence researchers' intention to use Generative AI tools in their research activities?

Participants receive a collection of studies related to technology adoption, AI adoption, trust in AI, perceived risks of GenAI, and the use of GenAI in higher education and research.

Their task is to use GenAI to organise this literature, identify the main research streams, synthesise their findings, identify relevant gaps and tensions, and determine whether the literature provides sufficient support for the proposed hypotheses.

LITERATURE CORPUS

The case provides participants with approximately 10 academic studies, represented by their bibliographic information, title, and abstract.

The corpus is deliberately heterogeneous. The papers do not all directly investigate researchers' adoption of GenAI. Instead, they provide different theoretical and empirical perspectives that can contribute to explaining the phenomenon.

FIELD	DETAIL		
	Article title	Year	Abstract / article page
1	Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology — Davis, F. D.	1989	View article / abstract
2	User Acceptance of Information Technology: Toward a Unified View — Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D.	2003	View article / abstract
3	Analysing ChatGPT Adoption Drivers with the TOEK Framework — Jo, H., & Bang, Y.	2023	View abstract and article
4	Human Trust in Artificial Intelligence: Review of Empirical Research — Glikson, E., & Woolley, A. W.	2020	View abstract
5	Fabrication and Errors in the Bibliographic Citations Generated by ChatGPT — Walters, W. H., & Wilder, E. I.	2023	View abstract and article
6	Conceptualising AI Literacy: An Exploratory Review — Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S.	2021	View abstract and article
7	ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education — Kasneci, E., et al.	2023	View article / abstract
8	ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns — Sallam, M.	2023	View abstract and full article

FIELD	DETAIL		
9	ChatGPT in Research and Health Professions Education: Challenges, Opportunities, and Future Directions — Williams-Persad, A., & Sooknanan, P.	2024	View abstract
10	Academics as Adopters of Generative AI: An Application of Diffusion of Innovations Theory — Singh, S., & Strzelecki, A.	2026*	View abstract and article

PROPOSED HYPOTHESES

Participants receive four preliminary hypotheses developed by the fictional research team.

H1: Researchers' perceived usefulness of GenAI has a positive effect on their intention to use GenAI in their research activities.

H2: Researchers' trust in GenAI has a positive effect on their intention to use GenAI in their research activities.

H3: Researchers' perceived risk associated with GenAI has a negative effect on their intention to use GenAI in their research activities.

H4: Researchers' AI literacy has a positive effect on their intention to use GenAI in their research activities.

An important feature of the exercise is that participants should not assume that all four hypotheses are equally well supported. One hypothesis may have substantial direct empirical support, while another may rely mainly on indirect evidence.

PROPOSED ACTIVITIES

ACTIVITY 1 – HUMAN ANALYSIS OF THE LITERATURE CORPUS

Participants receive 10 academic studies and initially work without GenAI. They are asked to answer:

-Which broad themes can you identify in the literature?

-How would you initially organise the 10 studies into groups?

-Which studies appear most relevant to each research question?

-Which of the four proposed hypotheses appear to have the strongest literature support?

ACTIVITY 2 – USE GENAI TO IDENTIFY RESEARCH STREAMS

Participants upload or provide the titles and abstracts to a GenAI tool.

Participants then challenge this first classification by asking GenAI to organise the same literature according to different criteria.

Prompt:

Analyse the supplied academic literature and identify the main research streams relevant to understanding researchers' adoption of Generative AI. For each proposed stream, explain its defining characteristics and identify the studies belonging to it. Base your classification exclusively on the supplied literature.

ACTIVITY 3 – SYNTHESISE FINDINGS WITHIN AND ACROSS STREAMS

Participants ask GenAI to analyse each validated research stream.

Prompt:

For each of the validated research streams, synthesise the central concepts, main theoretical arguments, and empirical findings from the studies assigned to that stream. Identify both areas of consensus and contradictory findings. Do not introduce information that is not contained in the supplied abstracts. For every substantive conclusion, indicate which studies provide the supporting evidence.

Compare the research streams. Identify connections between them and explain how they provide complementary or competing explanations of researchers' intention to adopt GenAI.

ACTIVITY 4 – IDENTIFY GAPS AND TENSIONS

Participants ask GenAI:

Prompt:

Based exclusively on the supplied literature, identify potential gaps, unresolved questions, and tensions relevant to the research question: “What factors influence researchers' intention to use GenAI in their research activities?” For each proposed gap, explain which studies provide the basis for identifying it and indicate whether the gap can actually be demonstrated from this corpus.

ACTIVITY 5 – CONNECT THE LITERATURE TO THE HYPOTHESES

Participants provide GenAI with the four proposed hypotheses and ask it to map the literature onto them.

Prompt:

Evaluate each of the four proposed hypotheses against the supplied literature. For each hypothesis, distinguish between: (1) studies providing direct empirical support; (2) studies providing indirect or theoretical support; (3) studies providing contradictory or qualifying evidence; and (4) studies that appear relevant but do not provide sufficient evidence to support the hypothesised relationship. Do not treat conceptual relevance as empirical evidence.

Participants then answer three critical questions:

-Which hypothesis has the strongest evidence base?

-Which hypothesis requires additional literature before it can be adequately justified?

-Has GenAI incorrectly interpreted correlation, conceptual relevance, or discussion of a variable as evidence supporting a causal hypothesis?

About this resource

TITLE	From literature mapping to hypothesis development with generative AI
TOOLKIT REFERENCE	UC1-01, use case 1: Literature review and synthesis
DEVELOPED BY	UPF Barcelona School of Management, Ramon Bastida
SUGGESTED CITATION	AI Researchers (2026). From literature mapping to hypothesis development with generative AI. AI Readiness Toolkit, UC1-01. UPF Barcelona School of Management, Ramon Bastida. airesearchers.eu
QUALITY ASSURANCE	Reviewed by the AI Researchers consortium for quality, categorisation and alignment with the AI Engagement Framework. External peer review by senior academics is under way, and the item may be updated as a result.
VERSION	1.0, November 2026
PROJECT	AI Researchers: Embedding Responsible AI in Research. Erasmus+ Cooperation Partnership in Higher Education, KA220-HED-66835D2B. October 2025 to March 2028.
CONSORTIUM	University of Lodz (coordinator), FEP University of Porto, UPF Barcelona School of Management, ACEEU, NIPI (OPI PIB), European E-learning Institute and Momentum Marketing Services

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

AI Researchers
Embedding Responsible AI in Research