

THE ECOSYSTEM OF EVIDENCE

Navigating the future
of the ecosystem of evidence

Artificial intelligence and digital transformation
in research and healthcare

11th EBHC International Conference
Taormina, 21st – 24th October 2026

#EBHC2026

Human-Centered AI for Evidence Synthesis:

A Scoping Review of User Experience (UX) & ~~Adoption~~

Chelsea Valenzuela
PhD Candidate
JBI, Adelaide ~~University~~
Australia

Valenzuela, Chelsea; Mine, Koya PhD; Hasanoff, Sabira; Holmes, Zachary PhD; Arno, Anneliese PhD; Stern, Cindy PhD; Aromataris, Edoardo PhD

30th 1996-2026 GIMBE EVIDENCE FOR HEALTH

ACKNOWLEDGEMENTS

- **Co-authors:**
  - Dr Koya Mine, Sabira Hasanoff, and Dr Zachary Holmes for their contributions to this review
  - **My PhD supervisors:** Associate Professor Edoardo Aromataris, Dr Cindy Stern, and Dr Anneliese Arno for their contributions, feedback, and support

- **Additional acknowledgements:**
  - Vikki Langton, Adelaide University, for assisting with the development of the search strategy
  - Justin Clark, Bond University, for providing guidance and feedback early on in my PhD

- This work will contribute toward a doctor of philosophy (PhD) degree from Adelaide University

Background

- Automation/AI tools for evidence synthesis are growing rapidly
- Accuracy is important, but that alone does not drive uptake: UX, trust and workflow fit also matter
- No review has mapped how UX and adoption have been evaluated and operationalised

Aims & Research Q's

AIM: To map the evidence on user experience (UX) and adoption of automation/AI for evidence synthesis

RESEARCH QUESTIONS

- How have UX and adoption of automation and AI been evaluated?
- What UX principles and corresponding design features are discussed?

Methods

Scan QR Code for
Scoping Review Protocol
DOI: 10.11124/JBIES-25-00236

- **JBI methodology, PRISMA-ScR reporting**
- **Inclusion criteria:** Concept (UX and adoption), Context (automation/AI technologies for ES types and stages)
- **Sources:** PubMed, Web of Science, Scopus, ProQuest, ACM Digital Library, and grey lit.
  - From 2015, no language limits
  - Initial search June 2025 (presented today); update search June 2026 (in screening)
- **Screening:** dual-independent – Rayyan (Ti/Ab) and Covidence (FT)
- **Extraction:** piloted subset in duplicate; duplicate extraction in progress
- **Deductive-inductive analysis:** UX principles, human-AI interaction factors, technology adoption frameworks

Included Studies

140 studies included (between 2015 to 2026)

40 empirically evaluated UX, adoption or both

100 discussed UX or adoption only

Levels of technology

Platform
e.g., Rayyan, Covidence, JBI SUMARI

Tool *(level most discussed)*
e.g., Rayyan screening tool

Feature
e.g., Rayyan relevance ranking

Automation to AI spectrum

More predictable

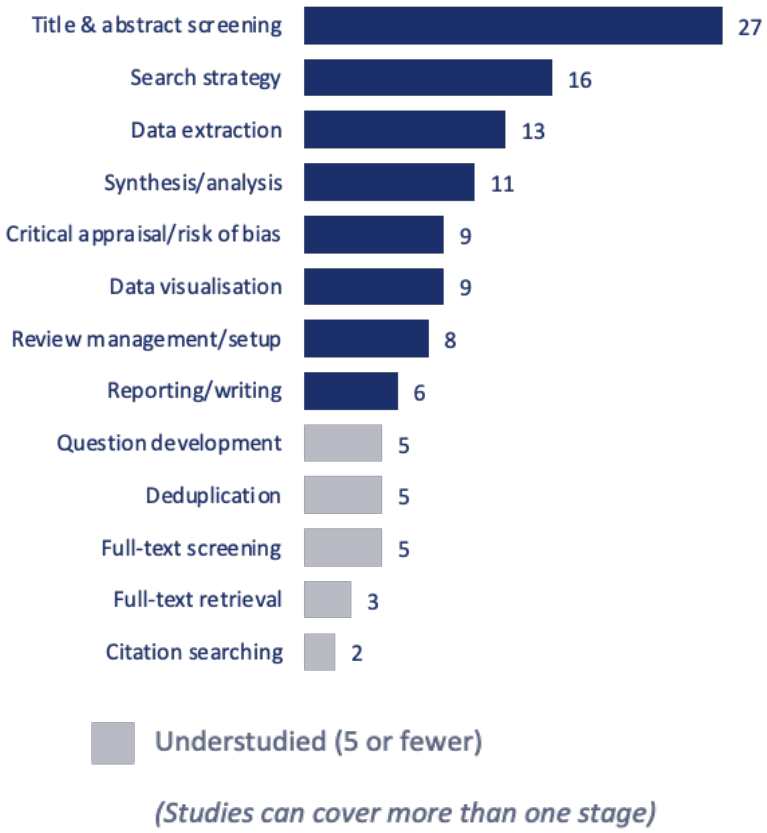
Rules-based automation
Fixed rules, no learning

Machine learning (ML)
Learns from data and patterns, incl. NLP and text mining

Generative AI / LLMs
Creates new content

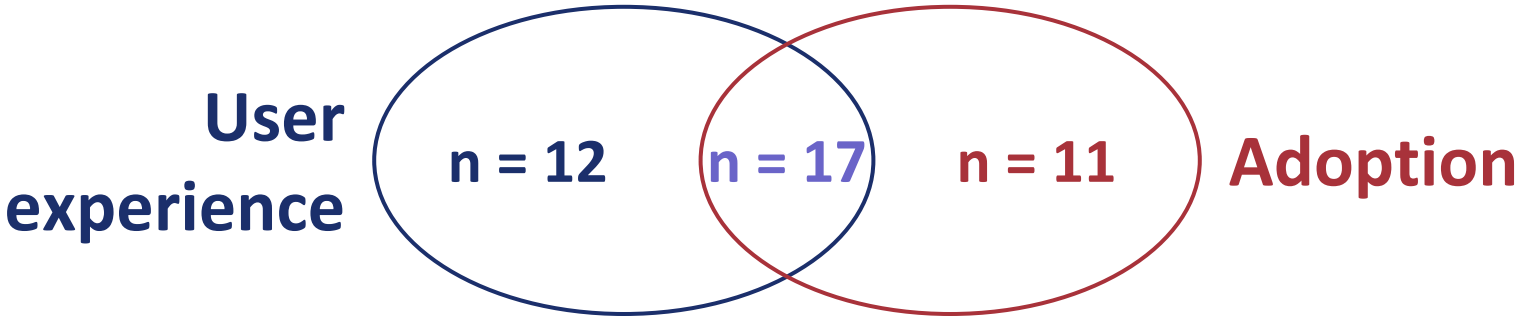
Trust concerns centred on ML and generative AI: accuracy, transparency, user control

Evidence synthesis stages supported



(Studies can cover more than one stage)

UX/Adoption Evaluations



UX only

WHO EVALUATED

Developers testing their own tools, and independent expert evaluators

FOCUS

The interface of specific tools in screening and searching

HOW

Usability instruments and user feedback surveys

Common terms: user-friendly, easy to use

Both (largest group)

WHO EVALUATED

Widest range: developers, independent experts, and recruited external cohorts

FOCUS

Tools tested with real users, from randomised user studies to live reviews

HOW

Performance metrics plus usability measures

Common terms: usability, workflow fit

Adoption only

WHO EVALUATED

Recruited and independent expert evaluators

FOCUS

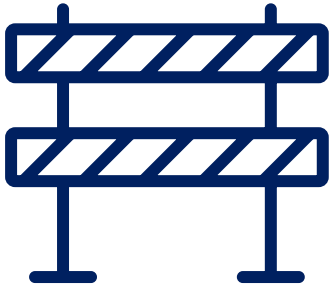
Acceptability, workflow fit and tolerance of error

HOW

Expert walkthroughs, surveys and qualitative sessions

Common terms: uptake, acceptability, intention to use

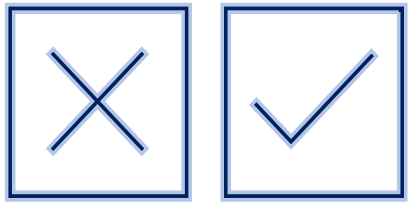
UX Principles & Design Features



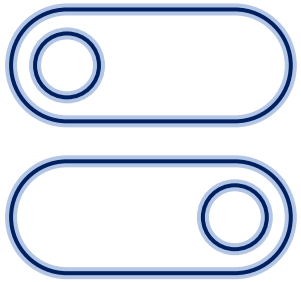
Constraints



Conceptual Model



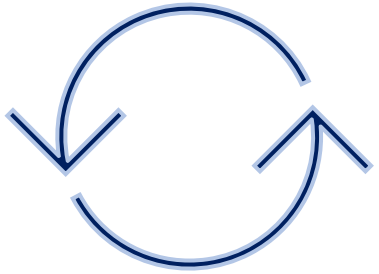
Signifiers



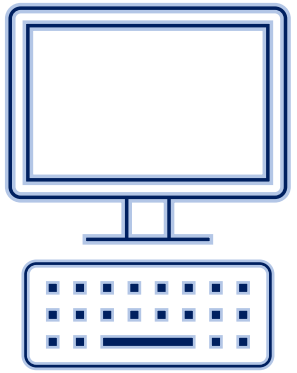
Affordances



Discoverability



Feedback



Mapping

Norman DA. The design of everyday things. Revised and expanded ed. Cambridge, MA: The MIT Press; 2013.

Limits

This review

- Preliminary findings from the initial search; updated search in screening
- Varied terminology required interpretive coding

The evidence

- Mostly short-term usability or intention to use, not sustained adoption
- Variable methods and reporting
- Rapidly evolving field

Conclusions

- In addition to accuracy, user experience is an important factor for adoption, but it isn't yet evaluated consistently or over time
- AI technologies that keep humans in control, explain their outputs and fit familiar workflows are more likely to be trusted
- Next steps: longer-term, co-designed evaluations across evidence synthesis roles

Thank You ☺

chelsea.valenzuela@adelaide.edu.au