The Literature Review Seminar

Tools

- Distinguish the major approaches of setting up tools for literature reviews
- Practice the use of an open-synthesis platform (CoLRev)
- Appreciate how AI and genAI/LLM may change the conduct of literature reviews

Start the demo

Start the demo (account and login on GitHub required)

Typical setups

Overall, there are many tools for literature reviews. The systematicreviewtoolbox.com alone listed over 250 tools (**discontinued in early 2024**).

There are three major approaches:

- **Self-managed approach**: Combine multiple tools, including a reference manager, and Excel
- **Platform**: Select a platform that handles the whole workflow and use the default functionality or extensions

Self-managed approach

Common elements:

- **Reference manager** to import, deduplicate, screen, extract data, analyze, and cite search results (e.g., Zotero, Endnote, Citavi, Mendeley, Jabref)
- **Excel** can be used for the screen, data extraction, and analysis
- **Specialized tools** for individual steps (see next slide)
- **Word processor** for write-up

Self-managed approach: Tools

Leading automation tools for literature reviews (Wagner et al. 2021):

Step	Research Tools
Search	*LitSonar*: Supports search query translation.
	litsearchr: Supports search strategy development.
	connectedpapers, inciteful: Support citation searches.
	TheoryOn: Supports construct searches.
Screen	*ASReview*: AI-supported screening.
Quality Assessment	*Robot Reviewer*: AI-supported quality appraisal.
Data Analysis	*Obsidian*: A tool for knowledge management and data extraction.
	RevMan: A tool to conduct meta-analyses.

Self-managed approach

Advantages:

- Low cost and quick setup
- Relatively high flexibility to use different tools and pursue different goals (review types)

Disadvantages:

- Data is handled manually: assigning IDs, sharing PDFs, keeping track of the status of records, data conversion, manual import and export
- Error-prone, especially when using Excel (see 1, 2)
- Individual tools may have limited compatibility
- Working in a team requires explicit and careful coordination
- Updating searches is challenging

Platforms

Criteria	CoLRev	LitStudy	BUHOS	Covidence
Review types				
Supports different genres of review methods	✓	✗	✗	✗
Process steps				
Review objectives and protocol	✓	✓	✓	✓
Search	✓	✓	✓	✓
Duplicate handling	✓	✗	●	●
(Pre)Screen	✓	●	✓	✓
Data extraction	✓	●	✓	✓
Data analysis and quality appraisal	✓	●	✓	✓
Synthesis and reporting	✓	✓	✓	✓

Platforms: CoLRev and open synthesis

We may envision an open and extensible research platform supporting different types of literature reviews.

The following aspects deem relevant:

- Shared data structures and processes
- Open-Source license and extensibility through packages
- Transparent data management, enabling the collaboration of reviewers and algorithms, including Artificial Intelligence and Generative Artificial Intelligence
- Self-explanatory workflow

Disclaimer: I am the lead developer of CoLRev.

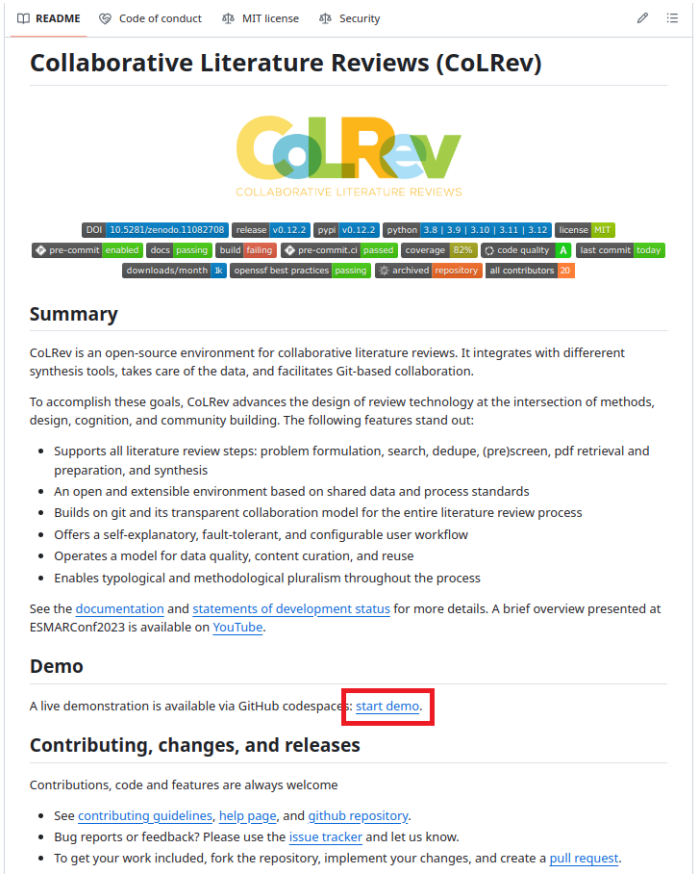
Platforms: CoLRev and open synthesis

- An open platform supporting all steps (see table below and demo in the documentation)
- Based on Git for data versioning and collaboration
- Extensible, offering different packages, e.g., packages for different types of reviews (not just "systematic reviews")

Step	Operations
Problem formulation	`colrev init`
Metadata retrieval	`colrev search`, `colrev load`, `colrev prep`, `colrev dedupe`
Metadata prescreen	`colrev prescreen`
PDF retrieval	`colrev pdfs`
PDF screen	`colrev screen`
Data extraction and synthesis	`colrev data`

Platforms: CoLRev and open synthesis

- Start the demo (account and login on GitHub required)
- Form small groups of 2-3 people
- Complete the tutorial
- Consult the documentation whenever necessary



README Code of conduct MIT license Security

Collaborative Literature Reviews (CoLRev)

CoLRev
COLLABORATIVE LITERATURE REVIEWS

DOI 10.5281/zenodo.11082708 release v0.12.2 pypi v0.12.2 python 3.8 | 3.9 | 3.10 | 3.11 | 3.12 license MIT
pre-commit enabled docs passing build failing pre-commit.ci passed coverage 82% code quality A last commit today
downloads/month 1k opensssf best practices passing archived repository all contributors 20

Summary

CoLRev is an open-source environment for collaborative literature reviews. It integrates with differerent synthesis tools, takes care of the data, and facilitates Git-based collaboration.

To accomplish these goals, CoLRev advances the design of review technology at the intersection of methods, design, cognition, and community building. The following features stand out:

- Supports all literature review steps: problem formulation, search, dedupe, (pre)screen, pdf retrieval and preparation, and synthesis
- An open and extensible environment based on shared data and process standards
- Builds on git and its transparent collaboration model for the entire literature review process
- Offers a self-explanatory, fault-tolerant, and configurable user workflow
- Operates a model for data quality, content curation, and reuse
- Enables typological and methodological pluralism throughout the process

See the documentation and statements of development status for more details. A brief overview presented at ESMARConf2023 is available on YouTube.

Demo

A live demonstration is available via GitHub codespaces: start demo.

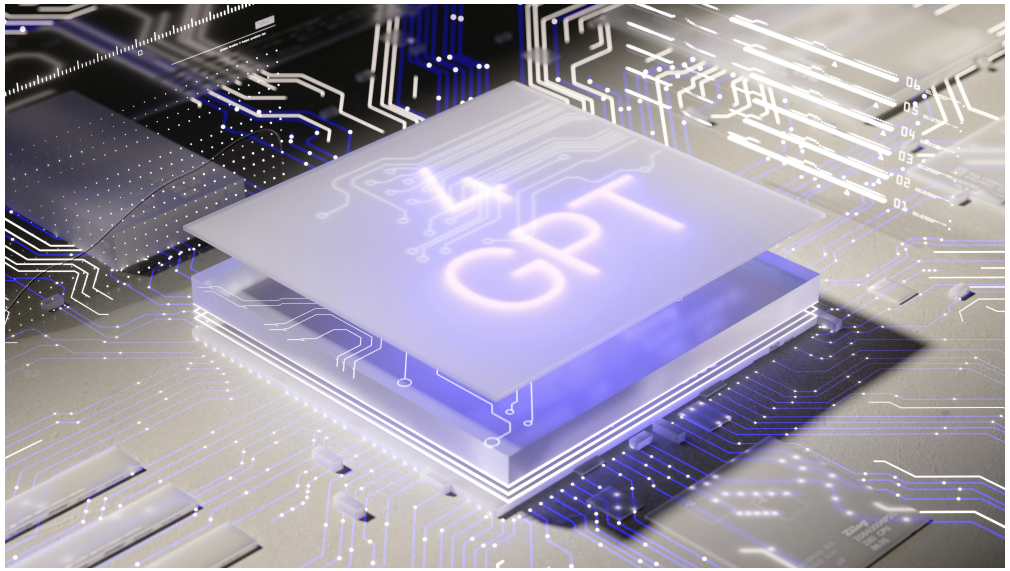
Contributing, changes, and releases

Contributions, code and features are always welcome

- See contributing guidelines, help page, and github repository.
- Bug reports or feedback? Please use the issue tracker and let us know.
- To get your work included, fork the repository, implement your changes, and create a pull request.

AI, genAI and the future(s) of literature reviews

Question: How would you use genAI-tools in a literature review?



LLMs, current challenges, and promises

Status quo: "Directly asking ChatGPT for research summaries does not produce compelling results"

- Language vs. knowledge and the problem of hallucination (fictitious references)
- LLMs do not necessarily have access to paywalled research
- Retrieval-augmented generation (APIs) as a potential remedy (e.g., consensus)

Researchers need to understand nuances of review types, methods, and steps

Which developments can be anticipated?

Review types

- Descriptive reviews may be the first to become obsolete given the summarizing capabilities of LLM
- For testing reviews, LLM can support different steps, including the generation of code for the analysis
- For reviews aimed at understanding or explaining, there may be different futures

Steps of the process

- LLM capabilities, or corresponding tools like litmaps, are particularly helpful for exploratory activities
- Language handling capabilities are useful for the design of queries in the systematic search phase (need to group synonyms)
- In the screen, restrictions of human cognitive capacities are one of the prime reasons to screen most of the papers based on the metadata (instead of the full-text). This could change with LLM, which
- Applications of LLM in the later steps have yet to be explored

Prompt example: Search query formulation

Best prompt identified by Wang et al. (2023):

```
You are an information specialist who develops Boolean queries for systematic reviews. You have extensive experience developing highly effective queries for searching the information systems literature. Your specialty is developing queries that retrieve as few irrelevant documents as possible and retrieve all relevant documents for your information needs. You are able to take an information need such as: “Review of IT Business Value” and generate valid Web of Science queries such as:
“TI=(IT OR IS OR …) AND TI=(value OR payoff OR …) AND TI=(firm OR business OR …)”.

Now you have your information needed to conduct research on “The effect of LLM on individual performance at work”, please generate a highly effective systematic review Boolean query for the information need.
```

⚠️ ChatGPT is useful for writing Boolean search queries in **high-precision reviews**, such as rapid reviews

Prompt example: Screen

Best prompt identified by Syriani et al. (2023):

```
Context: I am screening papers for a systematic literature review. The topic of the systematic review is the effect of generative AI on individual productivity for programmers. The study should focus exclusively on this topic.
Instruction: Decide if the article should be included or excluded from the systematic review. I give the title and abstract of the article as input. Only answer include or exclude. Be lenient. I prefer including papers by mistake rather than excluding them by mistake.

Task i:
•	Title: “Twelve tips to leverage AI for efficient and effective medical question generation”
•	Abstract: “Crafting quality assessment questions in medical education […]”
```

⚠️ Performance of LLM-based screening varies considerably across datasets, indicating **limited generalizability**

⚠️ The findings show that LLMs does not consistently perform better than random classification (in terms of recall)

Summary

- Carefully assemble your toolkit by considering the
  - Fit with the type of review
  - Need for collaboration in a team
  - Compatibility between tools (effort for data management and conversion)
- Consider open-synthesis platforms such as CoLRev
- Understand how AI and genAI/LLM may facilitate or change the process

We value your feedback and suggestions

We encourage you to share your feedback and suggestions on this slide deck:

 <img src="../assets/iconmonstr-pencil-lined.svg" alt="Edit" width="32" height="32"> Suggest specific changes by directly modifying the content
 <img src="../assets/iconmonstr-info-12.svg" alt="New Issue" width="32" height="32"> Provide feedback by submitting an issue
Your feedback plays a crucial role in helping us align with our core goals of **impact in research, teaching, and practice**. By contributing your suggestions, you help us further our commitment to **rigor**, **openness** and **participation**. Together, we can continuously enhance our work by contributing to **continuous learning** and collaboration across our community.

Visit this page to learn more about our goals: 🚀 🛠️ ♻️ 🙏 👨‍🎓 .

The next steps...

- Continue to develop the review protocol
- Schedule meetings to discuss the protocol as needed

Thank you!

- Thank you for participating in the seminar
- Keep in mind: If you work on literature reviews, there are opportunities to reconnect!
- Give us some feedback
- Help us spread the word to other students

References

Syriani, E., David, I., and Kumar, G. 2023. “Assessing the Ability of ChatGPT to Screen Articles for Systematic Reviews,” arXiv. doi:10.48550/ARXIV.2307.06464.

Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, 37(2), 209-226. doi:10.1177/02683962211048 20

Wang, S., Scells, H., Koopman, B., and Zuccon, G. 2023. “Can ChatGPT Write a Good Boolean Query for Systematic Review Literature Search?” in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1426–1436. doi:10.1145/3539618.3591703.