AI Literacy

Tutorial of the Philosophy and Psychology Library

AI Act

The EU AI Act, also known as the AI Act, is the **world's first comprehensive legislation on artificial intelligence**. It aims to regulate the development and use of AI within the EU to ensure that AI systems are **safe, transparent, non-discriminatory and environmentally friendly**. The law takes a risk-based approach that provides for strong regulation for risky technologies such as self-driving cars, while less risky technologies such as deposit return machines are less regulated. Educational institutions also have areas that count as high-risk systems, which means that universities must ensure that their AI systems are safe before they are used (EU AI Act, European Commission, 2023).

The usage of AI within university

The University of Vienna does **not have a general regulation** for the use of AI in education, which means that its use is **neither generally permitted nor prohibited**. It is important to emphasize that AI is not intended to replace students' skills and knowledge - examinations should be taken by the students! The decision on whether and how AI tools are used is the responsibility of the lecturers and examiners and is subject-specific. Transparency is particularly important here: lecturers should specify which tools are permitted before courses or examinations - a challenging task given the large number of tools. It must also be kept in mind that developments in this area are rapid and that current regulations can therefore change quickly (https://studieren.univie.ac.at/lernen-pruefen/ki-in-studium-und-lehre/).

Careful Handling of AI

Caution is required when using AI tools in educational contexts. Text generators such as ChatGPT can sometimes **provide incorrect information** and may reproduce **existing biases** from the training data. It is important not to rely solely on the linguistic quality of AI outputs, as this does not necessarily reflect intelligence. In addition, **most AI tools do not comply with the data protection regulations** and use user input as training data, which is why personal data should not be used. The quality of sources should generally always be **evaluated critically**, as many tools do not allow any pre-definition in terms of quality. It is advisable to use **different search strategies** and check information from different sources. AI tools can make work processes easier, but it is important to check the results and revise them if necessary.

The **ROBOT model** developed by Wheatley and Hervieux can help to assess the reliability, objectivity, Bias, Ownership and Type of the AI and evaluate the overall benefit of the AI. Ultimately, the users themselves are responsible for the results.

Ethical aspects

The **ethical aspects** associated with AI are of great importance (Coeckelbergh, 2020) and relate primarily to **security, data protection, discrimination and responsibility** (Hagendorff, 2020; Jobin et al., 2019). The design and regulation of AI technologies are controversial. Overarching ethical challenges **include the change in human self-image and world relations** through AI systems as well as **their influence on human self-image and communication**. The use of AI in higher education has an impact on the **understanding of humanity and education**, causing educational institutions to rethink their meaning and purpose in relation to AI. Dealing with generative AI, such as ChatGPT, requires a **clarification of educational goals at institutional and subject level**, with teachers and learners working together to find answers to these questions.

Is AI neutral?

The **training data for AI is often biased**. One example is ChatGPT, which in one analysis showed a left-leaning political orientation by asking the model to provide answers to questions from a political orientation test (Rozado, 2023). This underlines the fact that **AI models tend not to** be neutral and that the **goals of the developers** are incorporated into their functioning. The neutrality of an AI depends on its **development, training and deployment** and requires transparency in these processes. The predictions of such systems must therefore be viewed critically.

Should AI output be cited?

When dealing with AI results, e.g. ChatGPT outputs, it should be noted that **they** are **not retrievable by readers**, as no answer is displayed twice. Therefore, they must be treated as **algorithm outputs and cited accordingly**. It is important to indicate in a paper **how the tools were used** - this should be noted accordingly at the beginning of the paper. Another challenge is which sources the answers refer to, as **ChatGPT** is **not the primary source** for the information. Further information on exact citation methods can be found, for example, on the website of the Central European University.

What is AI literacy?

AI literacy encompasses the ability to critically evaluate AI technologies, to communicate and collaborate effectively with AI and to use AI as a tool online, at home and at work (Long, &

Magerko, 2020). A person with AI literacy is therefore able to critically scrutinize AI technologies and use them effectively in different areas of life.

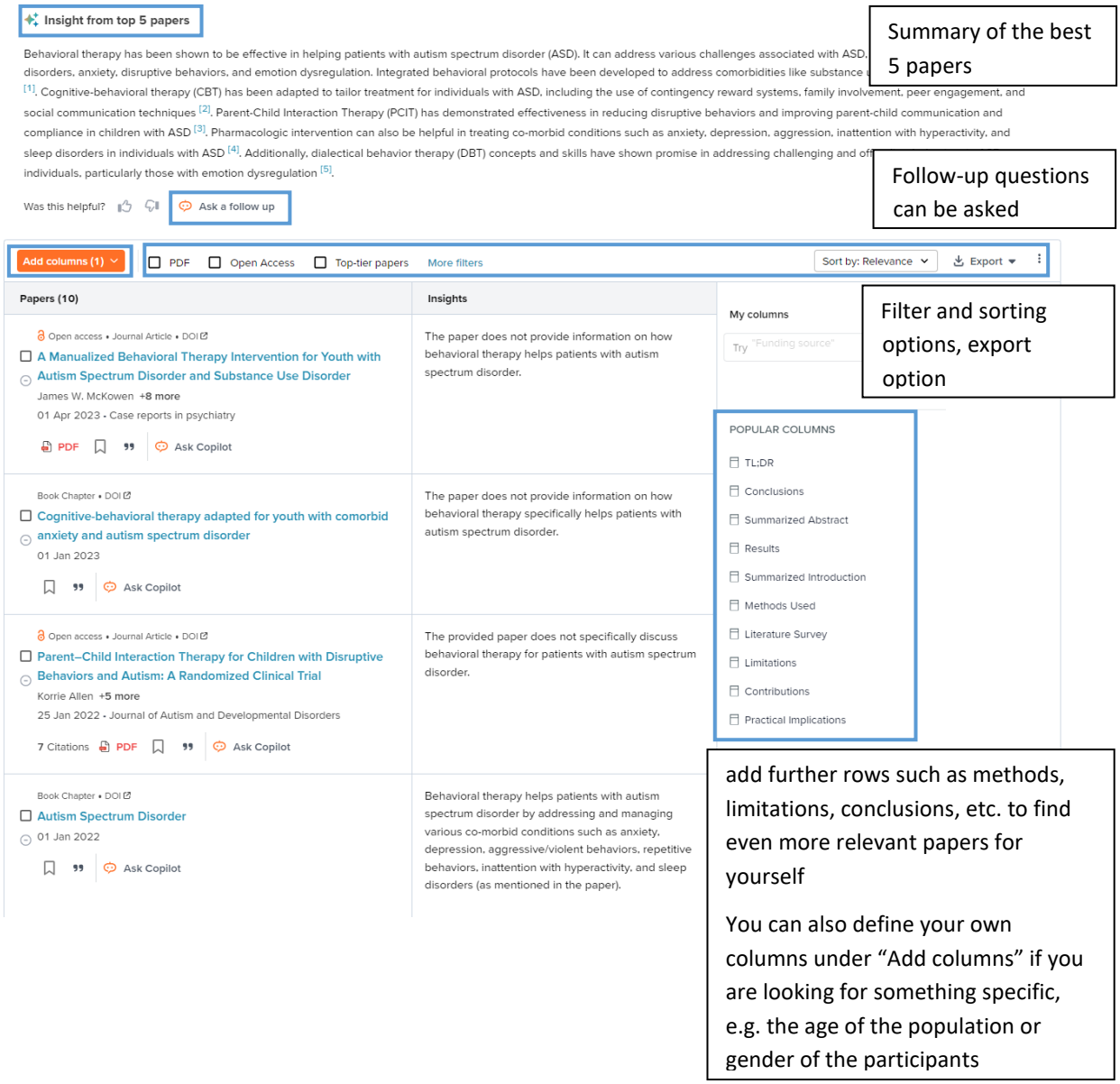
Selected AI tools

Important: The **tools are subject to ongoing development**, which means that changes to the user interface, functions, price structure and terms of use may occur at any time.

1. Scispace

- Link to Scispace
- Web-based tool with, simple structure
- Easy to navigate and use
- Does not require registration (but search is limited and searches cannot be saved)
- Free version: limited questions (no information on how many questions are for free)
- Full version currently costs 12 US dollars per month
- Offers additional tools such as extraction of data or a paraphraser
- Extension for Chrome available
- Can also be used in ChatGPT
- Currently only available in English
- Has its own comprehensive database of over 270 million scientific papers, authors, topics, journals and conferences

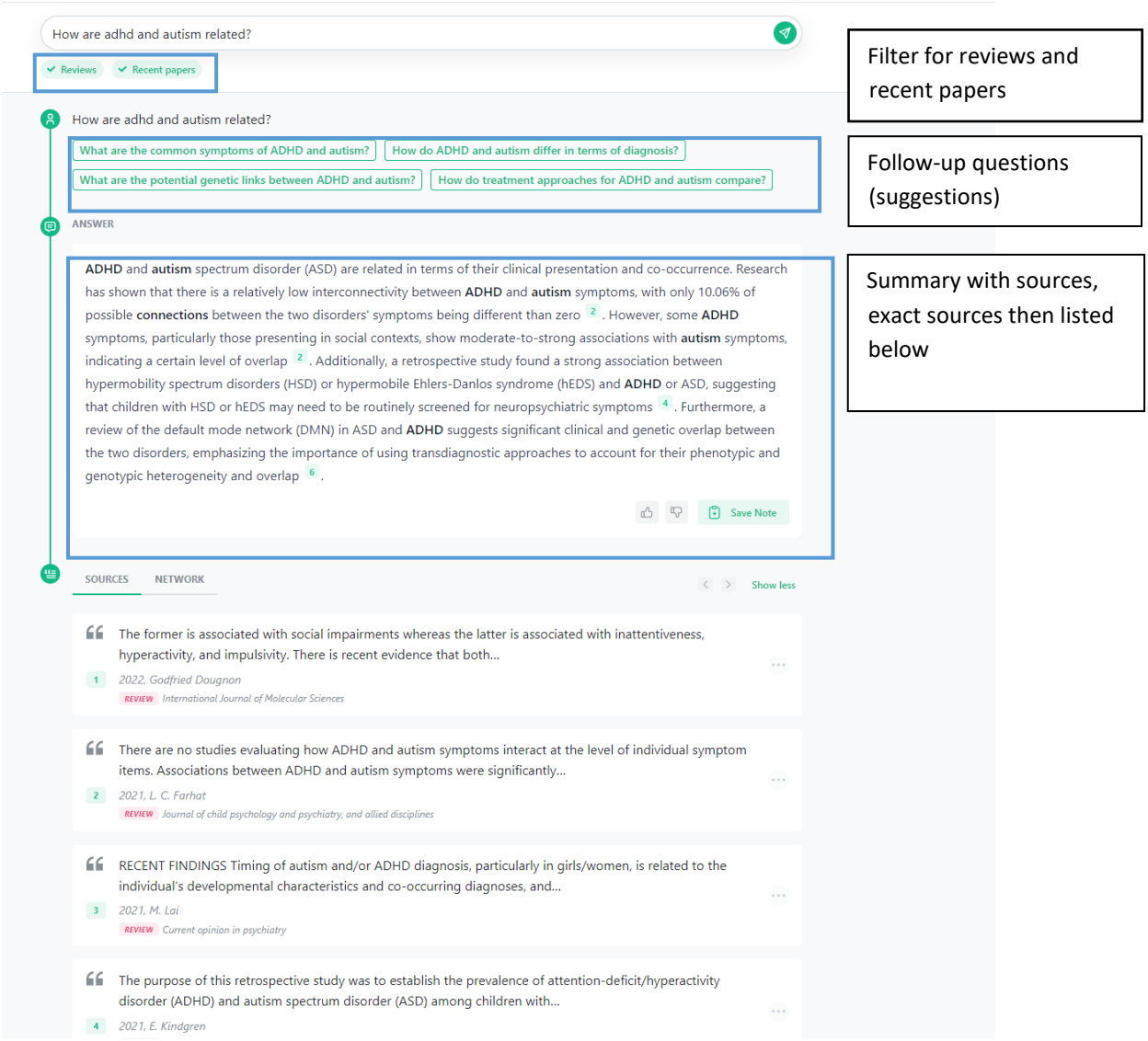
Example prompt in Scispace: „How does Behavioral Therapy help patients with autism spectrum disorder?"



2. ScienceOS

- Link zu ScienceOS
- web-based tool
- easy to navigate
- requires registration
- free and unlimited
- fewer functions than Scispace
- based on Semantic Scholar

Example prompt ScienceOS: „How are adhd and autism related?"



Filter for reviews and recent papers

Follow-up questions (suggestions)

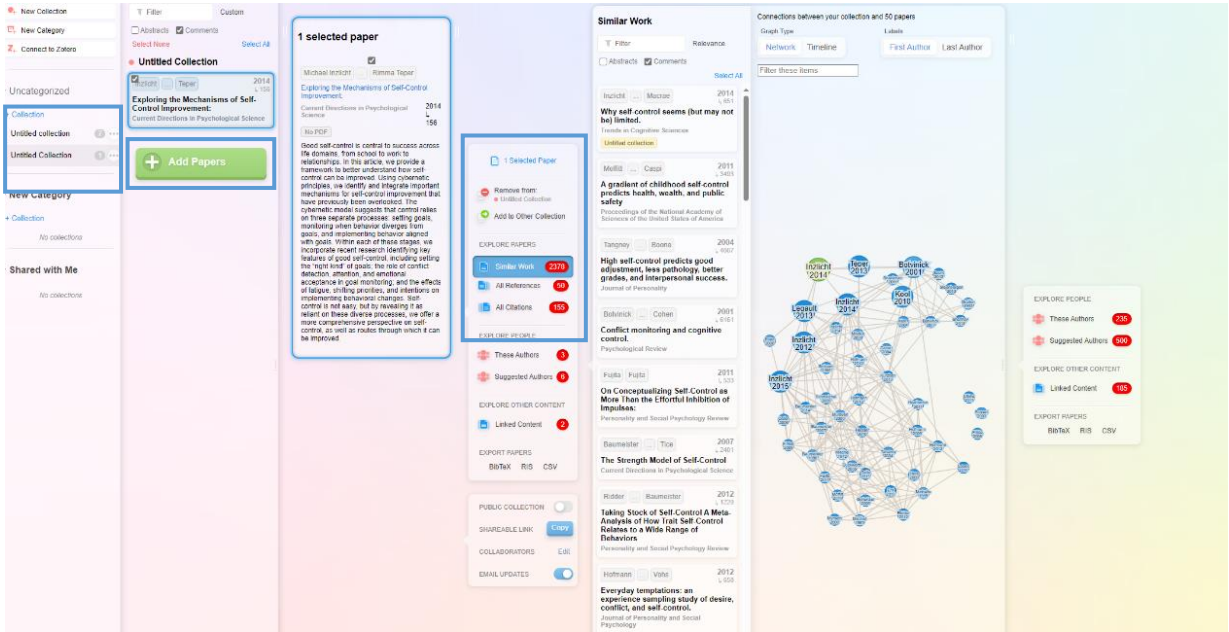
Summary with sources, exact sources then listed below

3. *ResearchRabbit*

- Link zu ResearchRabbit
- web-based tool
- easy to navigate
- requires registration
- free of charge and unlimited
- finds similar papers or papers that cite this work
- visualizes relations
- uses PubMed and Semantic Scholar
- Company claims that its unique database of "hundreds of millions of scientific articles" is the second largest database (after Google Scholar (ResearchRabbit - PMC (nih.gov))
- I.e.: access to databases such as PubMed as well as own database, not all sources fully transparent on website

Example ResearchRabbit

Create collections for easier organization	Add paper e.g. by DOI	Display and visualize similar works or works that cite this paper



4. *Consensus*

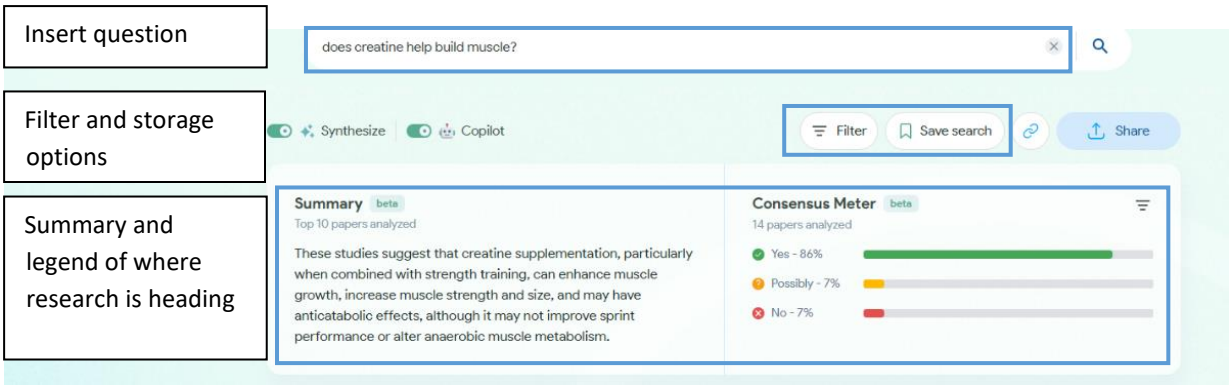
- Link zu Consensus
- web-based tool
- easy to navigate
- requires registration
- free of charge
- Entry of a research question, searches for papers that address it
- Legend showing where the research is heading
- Beta version, English only
- Data base: Semantic Scholar
- additionally uses Open AI's GPT-4 model to generate summaries of the results

Beispiel Consensus

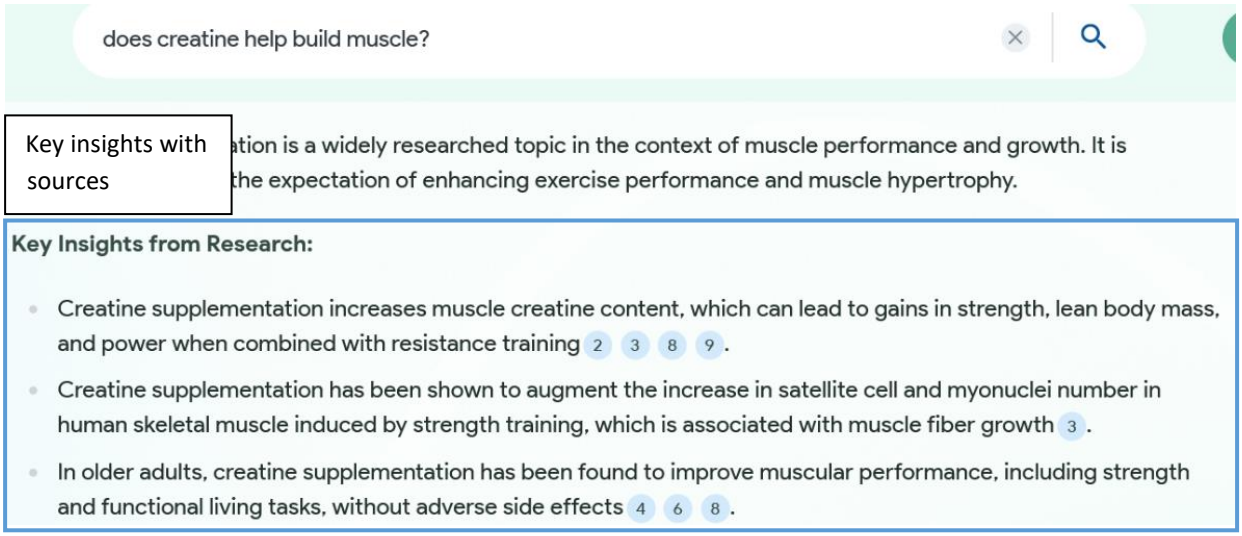
Insert question

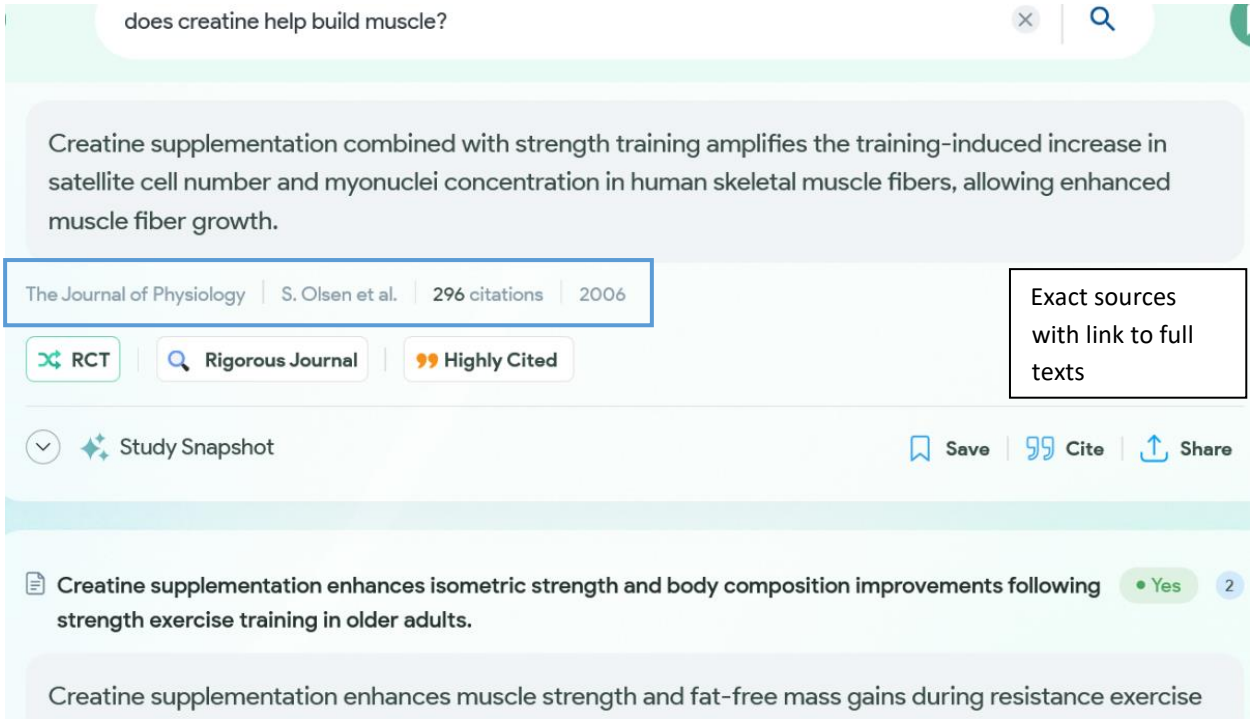
Filter and storage options

Summary and legend of where research is heading



Key insights with sources





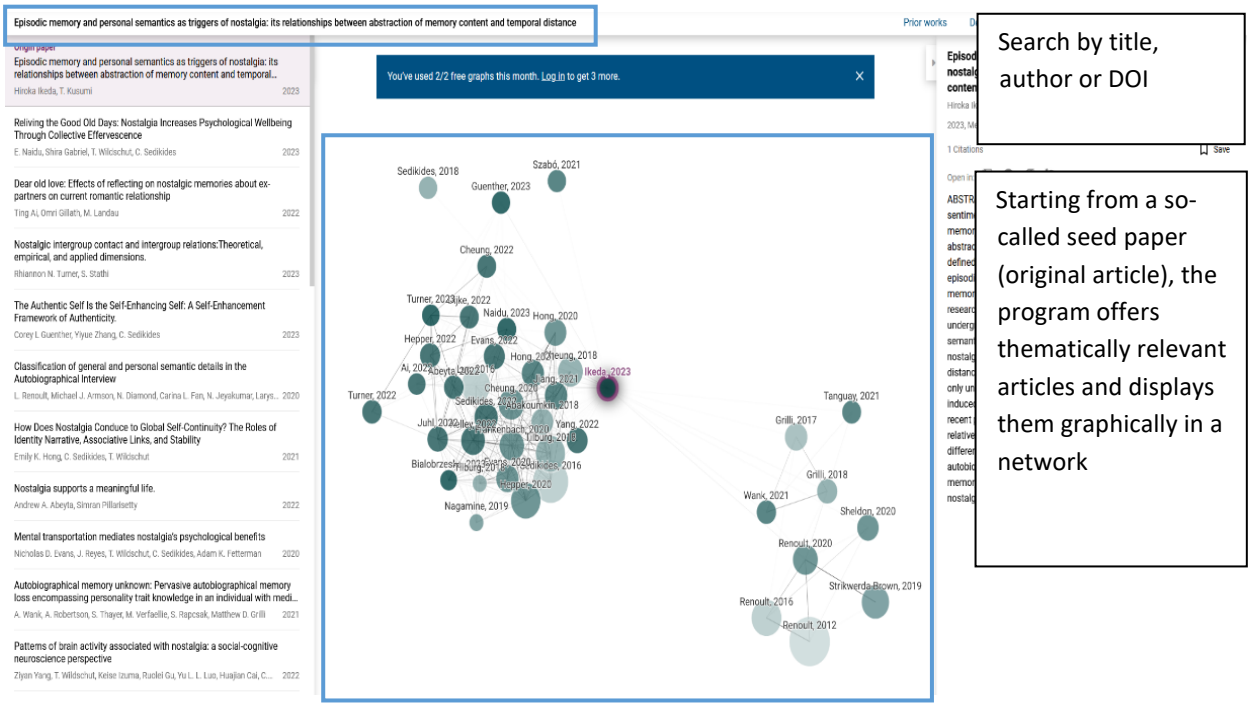
Literature Mapping Tools

Literature Mapping Tools are **software for visualizing bibliographic data** and a helpful tool for scientific work. They enable bibliometrics and the application of quantitative methods for measuring scientific performance. They are used to **provide an overview of research topics, find thematically similar articles, identify trends and gaps and refine search queries.**

1. Connected Papers

- Link zu Connected Papers
- web-based tool
- simple structure
- free version: 5 graphs per month
- Search based on a seed paper
- Data source: Semantic Scholar
- Export as Bibtex file

Example Connected Papers

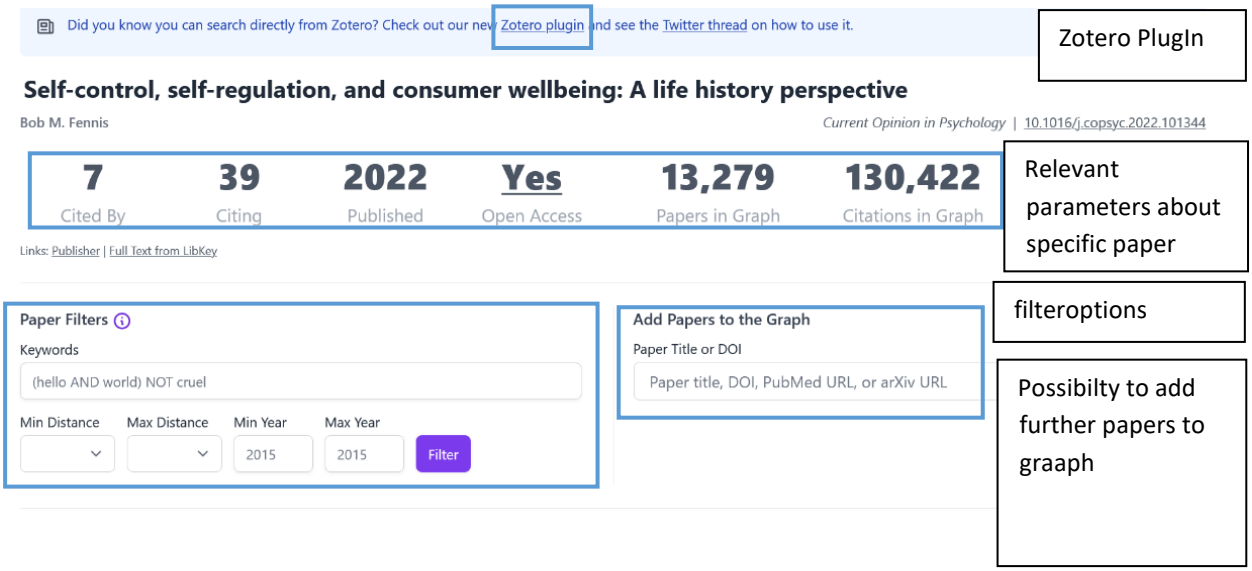


- the closer an article appears to the seed paper, the more thematically relevant it is
- the size of the nodes represents the number of citations and the color of the node represents the year of publication
- The Connected Papers algorithm calculates the similarity on the basis of co-citation and bibliographic coupling
- This means that new articles that have not yet been cited often can also be included in the search, as not only the analysis based on co-citations is used here, but also the bibliographic coupling
- This means that articles which do not cite each other but match due to their similarity can also be located next to each other (the graph is therefore not a tree diagram of citations).
- The exact procedure of the algorithm is not available publicly, but the program developers state that articles published in roughly the same generation are prioritized
- in addition to the graph, the program also generates a list of prior works = articles published earlier that were cited most frequently by the articles displayed
- derivative works = articles published later that cite several of the articles shown in the graph
- Articles can be exported as Bibtex files and imported into reference management programs

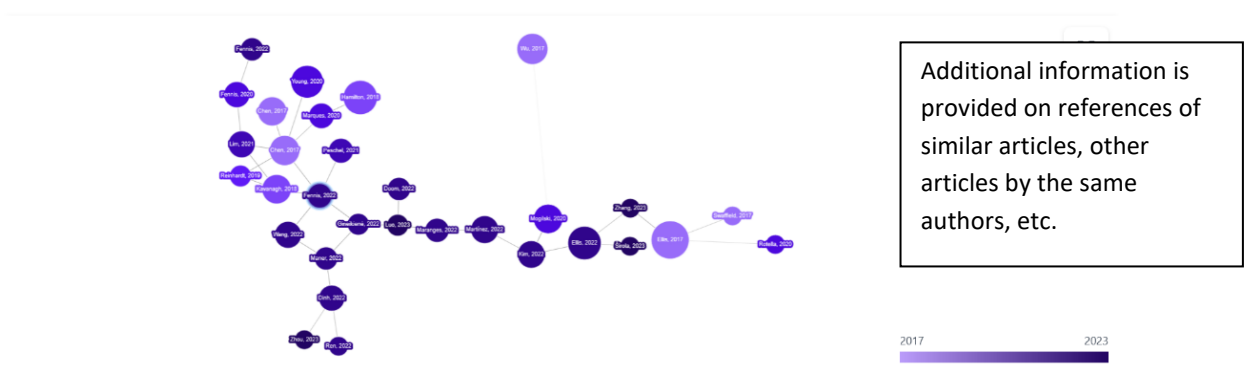
2. *Inciteful*

- Link to Inciteful
- free, web-based tool
- two core functions
  - Paper Discovery
  - Literature Connector
- Search based on a seed paper
- Recommendation: search with at least 5 articles
- Output of the program
  - similar papers
  - most important in the graph
  - recent papers by the top 100 authors
  - the most important recent papers
- Data basis: OpenAlex, SemanticScholar, CrossRef, OpenCitations

Example Inciteful



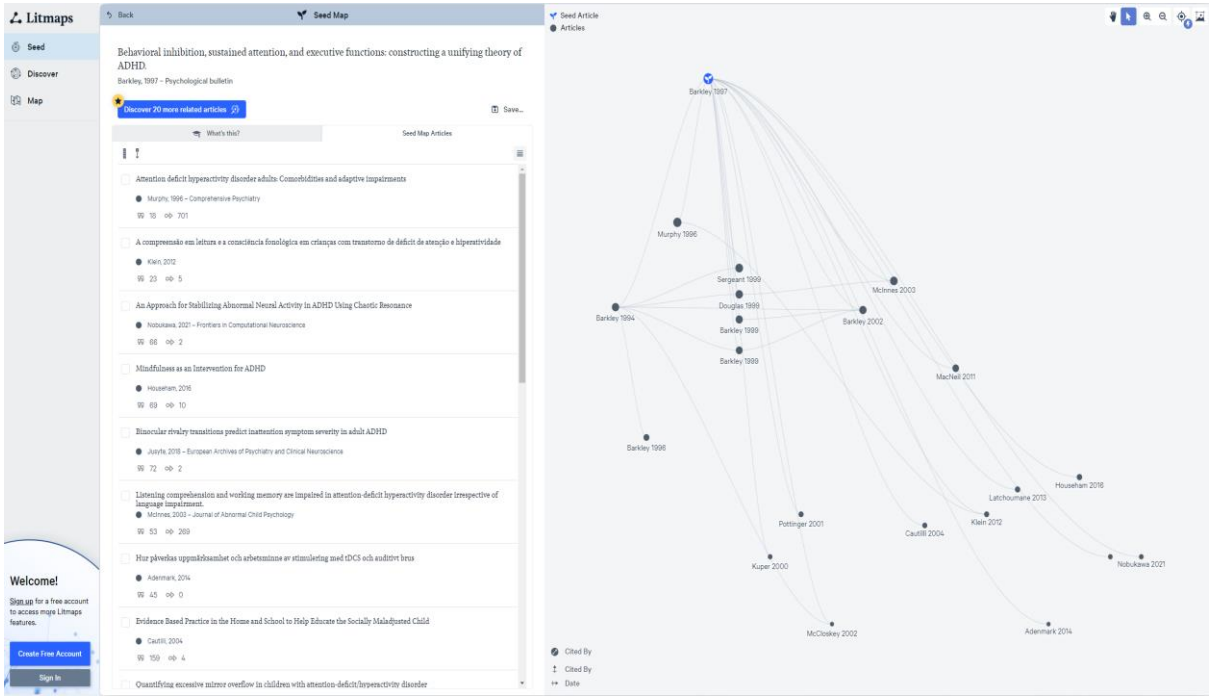
Visualisierung:



3. *Litmaps*

- Link to Litmaps
- web-based tool
- limited use within the free version
- strong focus on graphical representation of relationships
- various options for starting a search
- can display chronological relationships
- data sources are OpenAlex, Crossref, Semantic Scholar

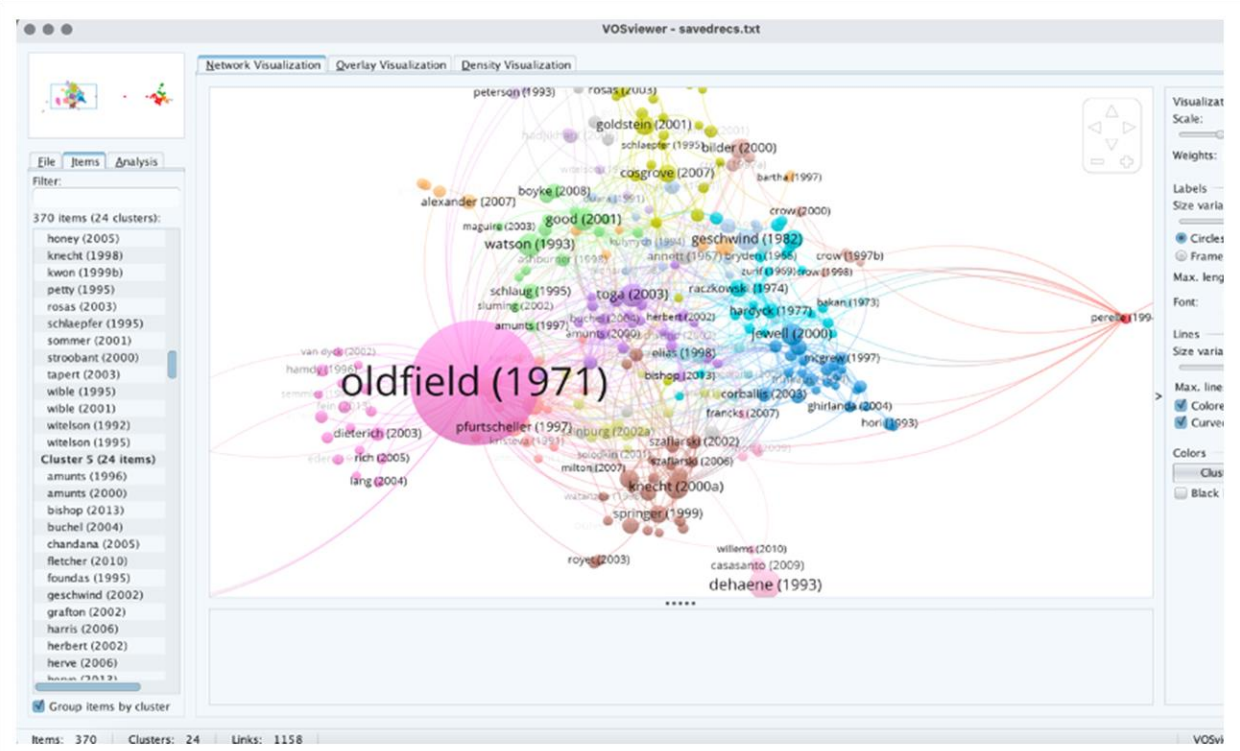
Example Litmaps



4. *VOSViewer*

- Link to VOSViewer
- free software
- desktop and web version
- advanced user requirements: independent data collection, cleansing and evaluation
- can generate networks fort he following data:
  - Keyword Co-Occurrence
  - Bibliographic Coupling
  - Co-Citation
- works with data from different sources
- Input possibilities are e.g. keyword searches in a database such as Web of Science. The search results are then exported to a text file and entered into the program
- Various sources from databases such as Web of Science, Scopus, Dimensions, Lens and PubMed to create networks such as co-authorship networks, citation-based networks and co-occurrence networks
- Crossref, Europe PMC and OpenAlex as well as Semantic Scholar, OpenCitations and WikiData are the base oft he networks

Example VOSViewer



Conclusion

The use of AI at the University of Vienna lies within **the responsibility of the lecturers and is subject-specific**, without general permission or prohibition. Student competencies are assessed with examinations, whereby permitted aids are determined by lecturers. A **critical examination of AI text generators** is essential, as results and evidence may be incorrect. It is important to observe **data protection regulations**, as users are responsible for AI results and tools should be continuously scrutinized. Various tools such as Scispace, ScienceOS, ResearchRabbit and Consensus are available for **literature searches** , while Connected Papers, Inciteful, Litmaps and VOSViewer, among others, can be used for **literature mapping**. These tools have **different data bases**, with Semantic Scholar, proprietary databases and collaborations being used particularly frequently.

Sources

Coeckelbergh, M. (2020). AI ethics. MIT Press. https://doi.org/10.7551/mitpress/12549.001.0001

Cole, V., & Boutet, M. (2023). ResearchRabbit. *Journal of the Canadian Health Libraries Association, 44*(2), 43-47. https://doi.org/10.29173%2Fjchla29699

Europäische Kommission. (2023). *A European approach to artificial intelligence*. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines, 30*(1), 99-120. https://doi.org/10.1007/s11023-020-09517-8

Hervieux, S. & Wheatley, A. (2020). The ROBOT test [Evaluation tool]. *The LibrAIry.* https://thelibrairy.wordpress.com/2020/03/11/the-robot-test

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Studienservice und Lehrwesen Universität Wien (n.d.). *KI in Studium und Lehre*. https://studieren.univie.ac.at/lernen-pruefen/ki-in-studium-und-lehre/

Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3313831.3376727

Pohlmann, P., Vossen, G., Everding, J., & Scheiper, J. (2022). Künstliche Intelligenz, Bias und Versicherungen – Eine technische und rechtliche Analyse. *Zeitschrift für die gesamte Versicherungswissenschaft, 111*(2), 135-175. https://doi.org/10.1007/s12297-022-00528-1

Rozado, D. (2023). The Political Biases of ChatGPT. *Social Sciences, 12*(3), 148. doi: https://doi.org/10.3390/socsci12030148