What should librarians know about artificial intelligence?

Milica Ševkušić

EIFL, Lithuania

Institute of Technical Sciences of SASA, Serbia

Image: Unsplash

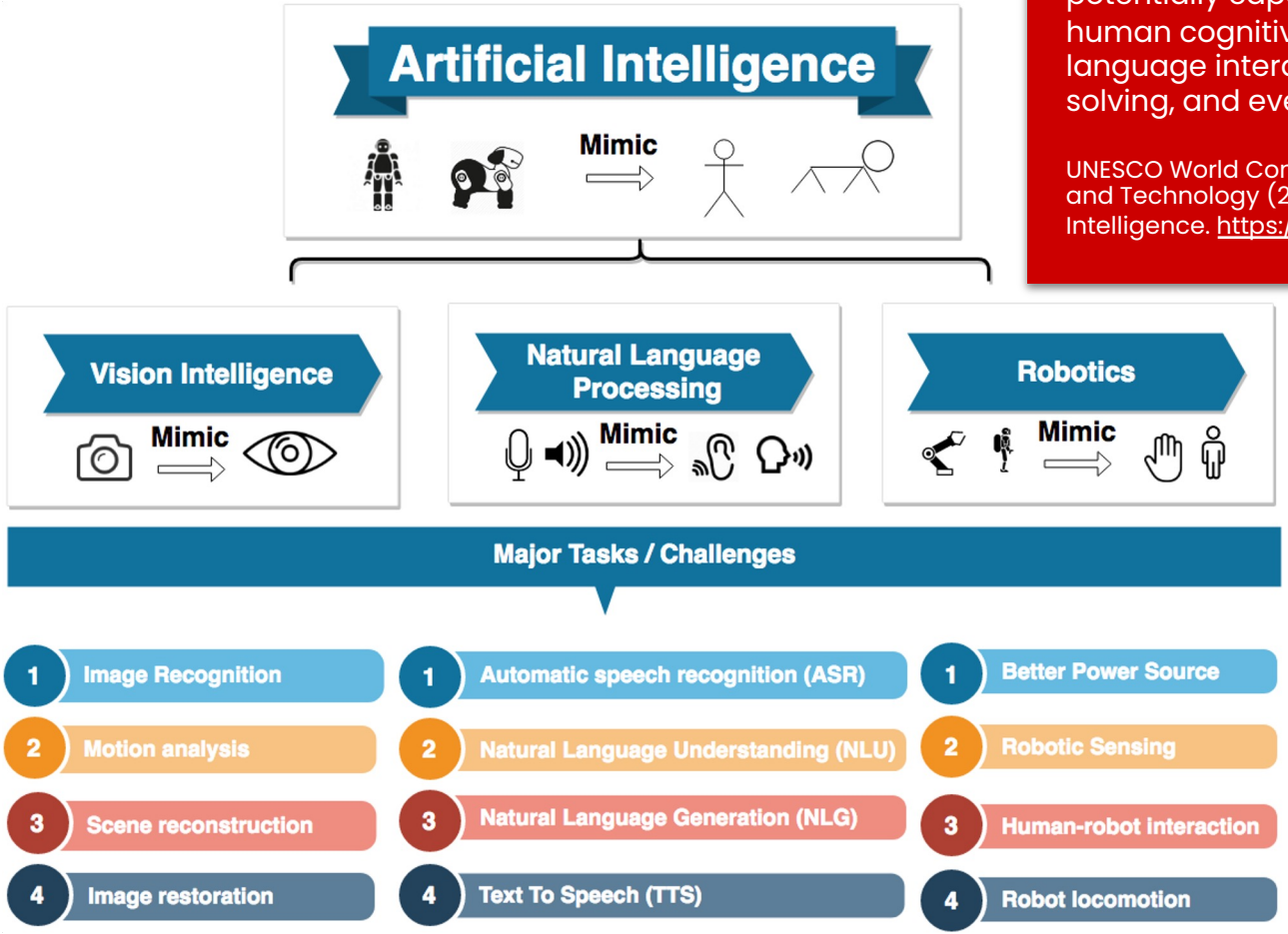
KLISC International Conference 2024, 22 May 2024

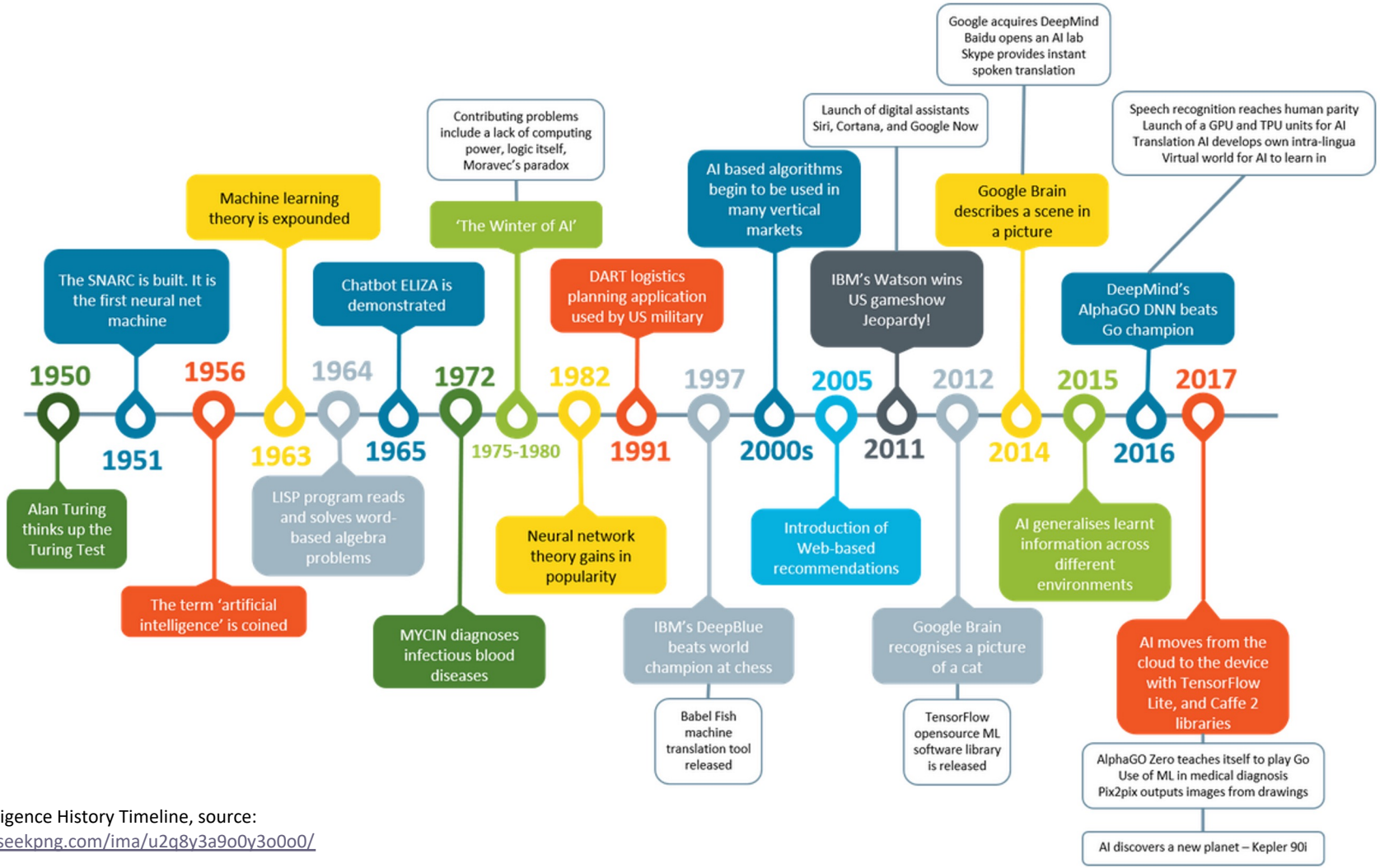
Outline

- Concepts
- AI and research libraries
  - AI literacy
  - AI, ethics and research integrity
  - AI and literature search
  - AI tools

There is no single or fixed definition of AI, but there is common agreement that machines based on AI "are potentially capable of imitating or even exceeding human cognitive capacities, including sensing, language interaction, reasoning and analysis, problem solving, and even creativity."

UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (2019). Preliminary Study on the Ethics of Artificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000367823





Artificial Intelligence History Timeline, source:
https://www.seekpng.com/ima/u2q8y3a9o0y3o0o0/

AI types based on capabilities

Narrow AI

Weak AI - the only type of AI that really exists today.

- can be trained to perform a single or narrow task (even to outperform a human)
- can't perform outside of its defined task

Virtual assistants: Siri, Amazon's Alexa, IBM Watson, Chatbots: OpenAI's ChatGPT

General AI

Artificial General Intelligence (AGI), Strong AI - a theoretical concept
If developed, it should be able to use previous learning and skills to accomplish new tasks in a different context without the need for human beings to train the underlying models. Could learn and perform any intellectual task that a human being can.

Super AI

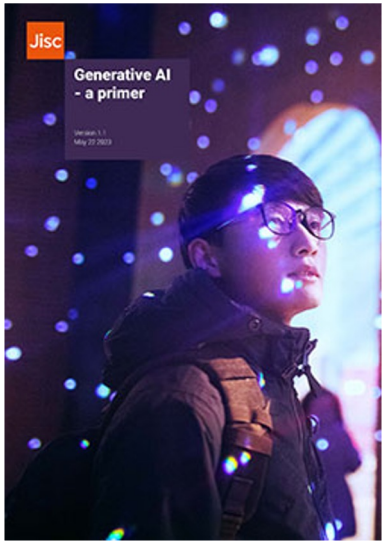
Super AI, artificial superintelligence and - a theoretical concept.
If ever realized, it would think, reason, learn, make judgements and have cognitive abilities that surpass those of human beings.

Generative AI

Deep-learning models that can generate text, images and other original content types based on the data they were trained on.

Large Language Models

- a subset of deep learning
- algorithm that can perform natural language processing tasks (recognize, translate, predict, generate text or other content
- can produce plausible but false information, often culturally or politically biased



‘Generative AI : A Primer’. 2023. JISC. https://beta.jisc.ac.uk/reports/generative-ai-a-primer.

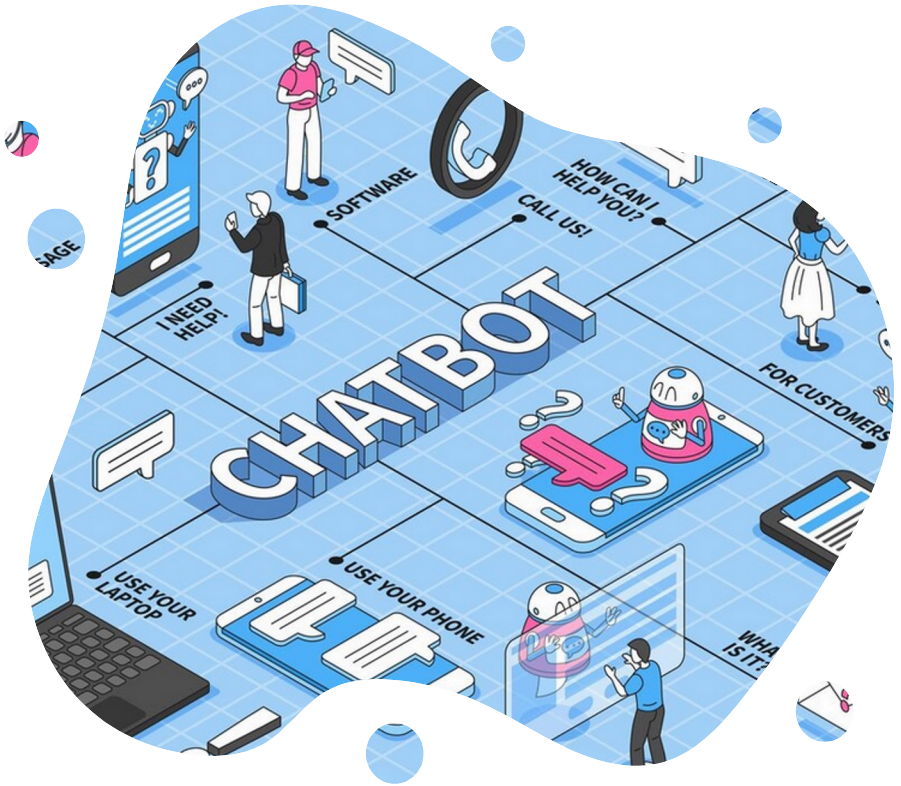
Chatbots

Chatbot - software application that aims to mimic human conversation (using text or voice)

- Natural language search
- Human language responses
- Personal assistance

Expected to help with:

- Text improvement
- Translation
- Text summarization
- Writing
- Literature reviews
- Visualization
- Coding



Generative AI and research libraries

- AI hype since late 2022 - even those not interested in AI are now asking questions
- Library users need support and guidance:
- What generative AI can and cannot do?
- Is it allowed and safe to use chatbots for research-related tasks?
- Ethical issues (plagiarism, fraud, fabrication, false accusations, etc.)
- How to embed AI in Research?

- Librarians are usually not AI/IT experts
- Information provided by the creators of AI tools is not always detailed and transparent
- The body of literature is growing, but how to find information quickly?
- Are library practices and workflows affected?
- How to embed AI in library workflows?

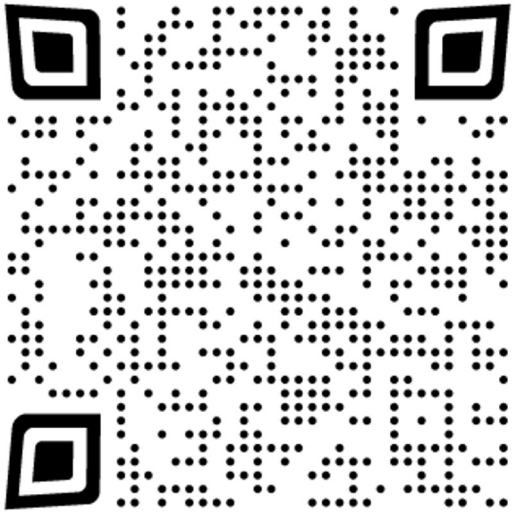
Learning resources

Training outline developed by EIFL: AI and Open Science

To be available on the EIFL website soon.

Module 5: AI and libraries

- Understanding challenges for libraries associated with AI
- Understanding the concept of AI literacy
- Being aware of new training topics relating to AI
- Being familiar with examples of new library services that can be built on AI



AI literacy

Associated with information literacy and digital literacy

A set of competencies that enables people to:

- critically evaluate AI technologies
- communicate and collaborate effectively with AI
- use AI as a tool both at home and in the workplace.
- the ability to understand the basics of how AI works, including machine learning, neural networks, and large language models
- the ability to use AI effectively and ethically
- the ability to make informed decisions about using AI technologies

TAKE ACTION!

- Know where to find information, reviews studies on AI.
- Be familiar with commonly used AI tools. Test them and read reviews.
- Be familiar with the main concepts and technologies and know the possibilities and limitations.
- Be familiar with regulations and recommendations.
- Instruct users how to check information.

'A Tech Librarian Explains How to Build AI Literacy'. 2024. *Choice 360* (blog). 24 April 2024. https://www.choice360.org/libtech-insight/a-tech-librarian-explains-how-to-build-ai-literacy/.

Ethics and research integrity

- Recommendations and best practice
- Using of generative AI in scholarly writings
- Privacy issues
- Copyright infringement in AI model training
- Copyright infringement arising from the use of AI tools

TAKE ACTION!

- Be familiar with regulations, best practice and recommendations.
- Inform researchers about new risks and remedies.
- Provide training.
- Put a strong focus on information and digital literacy!

Research integrity

SCIENCEINSIDER | SCIENTIFIC COMMUNITY

Science funding agencies say no to using AI for peer review

Concerns include confidentiality, accuracy, and "originality of thought"

14 JUL 2023 • 4:25 PM ET • BY JOCELYN KAISER

NEWS | 08 September 2023

Scientific sleuths spot dishonest ChatGPT use in papers

Manuscripts that don't disclose AI assistance are slipping past peer reviewers.

NEWS | 31 May 2023

AI intensifies fight against 'paper mills' that churn out fake research

Text- and image-generating tools present a new hurdle for efforts to tackle the growing number of fake papers making their way into the academic literature.

NEWS FEATURE | 10 October 2023

How ChatGPT and other AI tools could disrupt scientific publishing

A world of AI-assisted writing and reviewing might transform the nature of the scientific paper.

The Latest "Crisis" – Is the Research Literature Overrun with ChatGPT- and LLM-generated Articles?

Home › News › Technology › Software

ChatGPT used in peer reviews of Australian Research Council grant applications

[Submitted on 25 Mar 2024]

ChatGPT "contamination": estimating the prevalence of LLMs in the scholarly literature

Andrew Gray

Using AI to protect against AI image manipulation

"PhotoGuard," developed by MIT CSAIL researchers, prevents unauthorized image manipulation, safeguarding authenticity in the era of advanced generative models.

TECH IN YOUR LIFE

We tested a new ChatGPT-detector for teachers. It flagged an innocent student.

Five high school students helped our tech columnist test a ChatGPT detector coming from Turnitin to 2.1 million teachers. It missed enough to get someone in trouble.

Analysis by Geoffrey A. Fowler
Columnist | + Follow

Updated April 3, 2023 at 9:47 a.m. EDT | Published April 3, 2023 at 6:00 a.m. EDT

turnitin

Products · Solutions · Resources · Support · Contact Sales

Blog · 3 min read · Tuesday 4 April 2023

The launch of Turnitin's AI writing detector and the road ahead

FUTURE

How to Prove You Didn't Use ChatGPT: One Simple Trick to Avoid ChatGPT Plagiarism Accusations

By Amy D

Posted on May 22, 2023

https://www.turnitin.com/blog/the-launch-of-turnitins-ai-writing-detector-and-the-road-ahead

Recommendation on the Ethics of Artificial Intelligence

Paris, France
23 November 2021

Theme: Social & Human Sciences

Authoritative texts:
Arabic, Chinese, English, French, Russian, Spanish

WAME Revised Recommendations on Chatbots and Generative AI October 25, 2023

New revised WAME Recommendations: **Chatbots, Generative AI, and Scholarly Manuscripts: WAME Recommendations on Chatbots and Generative Artificial Intelligence in Relation to Scholarly Publications**

Searching

Lexical: combining keywords using logical operators; the search engine matches words;

Semantic: search by meaning; the search engine matches concepts;

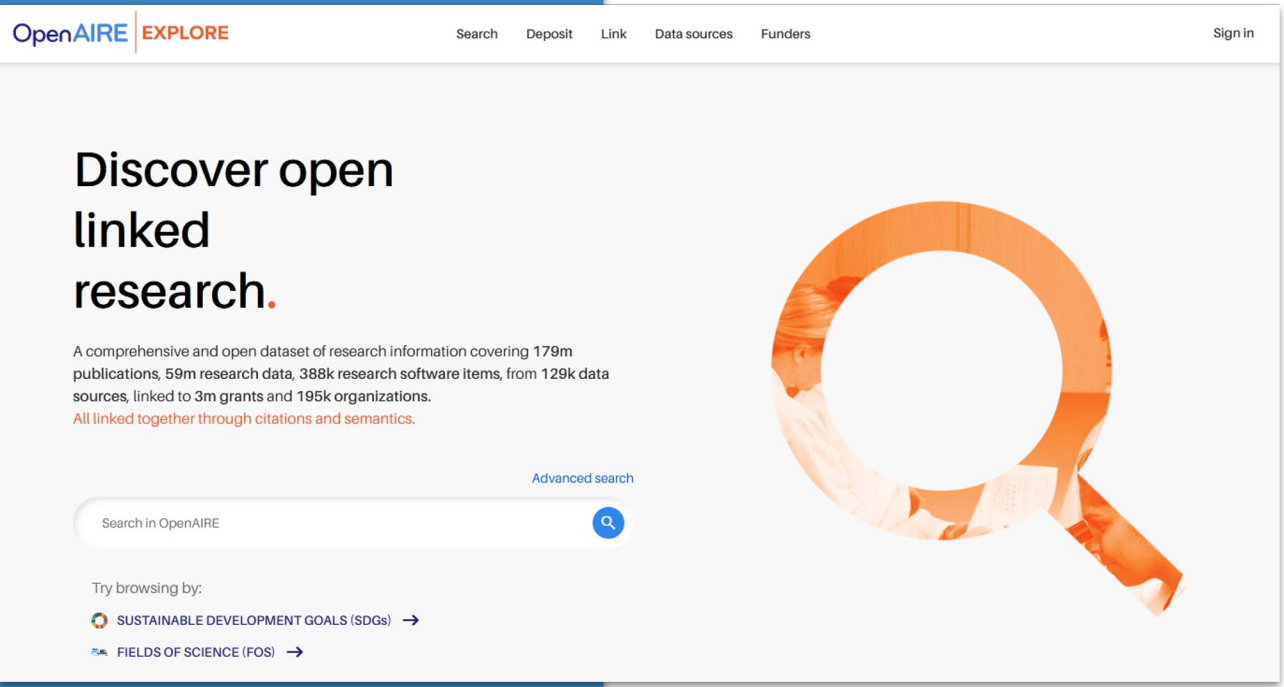
Natural language search: queries are in a natural language

AI technologies have been used for quite a while in various search engines

- What is indexed? (names and titles, abstracts, full text)
- How the data in a database are structured?
- What search strategies and search technologies are supported?
- How the search results are ranked? Is everything displayed or just a selection of results?
- Multilingual support

OpenAIRE Explore

https://explore.openaire.eu/



Natural Language Processing +

graph database +

graph mining

machine learning +

information retrieval +

Recommendation systems +

Visualization

Data source: OpenAIRE Research graph

Open Knowledge Maps

https://openknowledgemaps.org/

Natural Language Processing +

Graph database +

machine learning +

information retrieval +

Visualization

Data sources: BASE and PubMed

Information retrieval challenges related to Large Language Models (LLMs)

- Black box: the retrieval process is not transparent (which sources were taken into consideration, why and how a LLM arrived at a particular result)
- Searches are not reproducible
- Disputable interoperability
- The scope and quality of data used for training?
- LLMs (still) lack domain-specific knowledge (expensive to train, limited availability of data for training)
- LLMs are static: pre-trained, then frozen; have to be retrained
- Hallucinations
- Changing landscape

LLM + Retrieval-Augmented Generation (RAG)

- An AI framework for retrieving facts from external knowledge bases (very often Semantic Scholar) to supply large language models (LLMs) with up-to-date information
- Adds context
- Lowers the cost (by diminishing the need to retrain LLMs)

Tay, Aaron. n.d. “List of Academic Search Engines That Use Large Language Models for Generative Answers Using Retrieval Augmented Generation (RAG).” Aaron Tay’s Musings about Librarianship. Accessed November 16, 2023. https://musingsaboutlibrarianship.blogspot.com/p/list-of-academic-search-engines-that.html.

Name	Sources	LLM used	Upload your own PDF?	Produces literature review matrix?	Other features
Elicit.com/old.elicit.org	Semantic Scholar	OpenAI GPT models & other opensource LLMs	Yes	Yes	• List of concept search
Consensus	Semantic Scholar	GPT4 for summarises	No	No, has Consensus meter	
scite.ai assistant	Open Scholarly metadata and citation statements from selected partners	"We use a variety of Language models depending on situation." GPT3.5 (generally), GPT4 (enterprise client), Claude instant (fallback)	No	No	• Summaries include text from citation statements • Many options to control what is being cited
scispace	Unknown	Unknown	Yes	Yes	

Scopus AI: Change the way you view knowledge

Elsevier's new generative AI provides you with concise, trustworthy AI-based summaries of your queries.

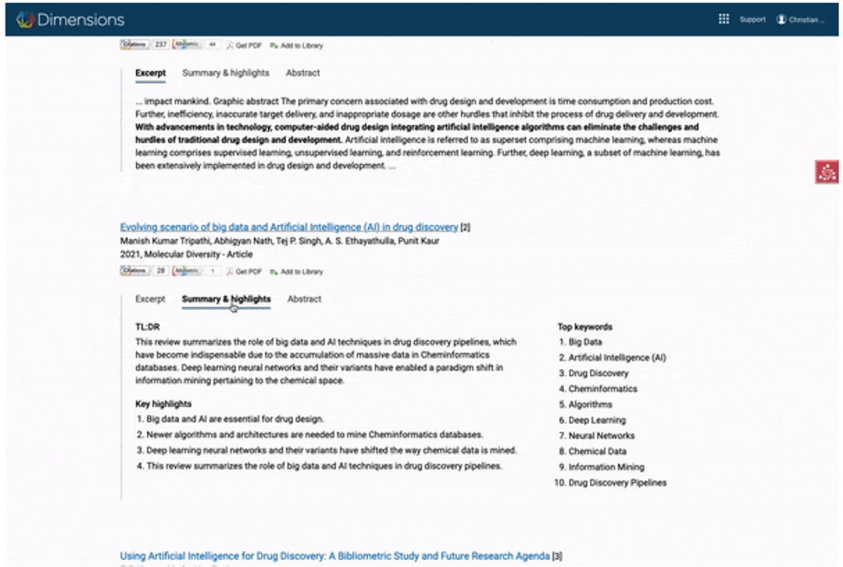
Whether you are looking to get up to speed on a new field quickly or find cross-disciplinary collaborators to push your research to the next level, Scopus AI can help.

CORE GPT (announced)

← ALL PRODUCTS

Dimensions AI Assistant

Taking data exploration to the next level




Elicit
Untitled notebook
New Notebook
Artificial Intelligence in Research
https://elicit.com/
The use of artificial intelligence in research
Summary of top 4 papers
Copy
The use of artificial intelligence (AI) in research is a rapidly evolving field with significant potential. Pigola (2023) notes that while AI is still in its early stages in academic research, its transformative impact is being recognized. Lara-Herrera (2024) emphasizes the need for ethical principles in the use of AI in education, particularly in thesis works and research projects. Tandalaskar (2024) provides an overview of AI's applications in various fields, including education, engineering, business, and medicine, highlighting its role in improving quality and efficiency. Waly (2024) explores the relationship between AI and scientific research, underscoring its ability to extract insights from large datasets, aid in predictive analytics, and streamline processes in laboratory settings. However, the integration of AI in research also presents challenges related to data quality, interpretability, and ethical considerations.
Add columns
Sort: Most relevant
Filters
Export as
PLUS

The us of artificial intelligence in research

Export References

The us of artificial intelligence in research

Artificial intelligence (AI) has become a transformative force across various fields of research. In the realm of healthcare, AI is being utilized for tasks such as aiding in the diagnosis of diseases like COVID-19 (Huang et al., 2021), improving intensive care unit practices (Nguyen et al., 2020), and even assisting in dermatology (De et al., 2020). The integration of AI in healthcare is driven by the potential to enhance medical decision-making processes and improve patient outcomes.

Moreover, AI is not limited to healthcare but extends to other sectors like the public sector (Wirtz et al., 2018), architecture (Harapan et al., 2021), and logistics (Zhu et al., 2022). In the public sector, AI applications are being explored to enhance operational efficiency and service delivery (Wirtz et al., 2018). In architecture, AI is revolutionizing design processes and enabling innovative solutions (Harapan et al., 2021). Additionally, in logistics, AI technologies are being harnessed to optimize modern automated systems (Zhu et al., 2022).

The evolution of AI research has led to the development of Explainable Artificial Intelligence (XAI) (Arrieta et al.,

References Search Strategy

"...Artificial intelligence (AI) is being used to aid in various aspects of the COVID-19 crisis, including epidemiology, molecular research and drug development, medical diagnosis and treatment, and socioeconomics...." >>

Section: Abstract

Artificial intelligence in the diagnosis of COVID-19: challenges and perspectives
Shigao Huang[1], Jie Yang[2], Simon Fong[3] et al. 2021 *Int. J. Biol. Sci.*

81 | 0 | 26 | 0

View full text Add to dashboard Cite

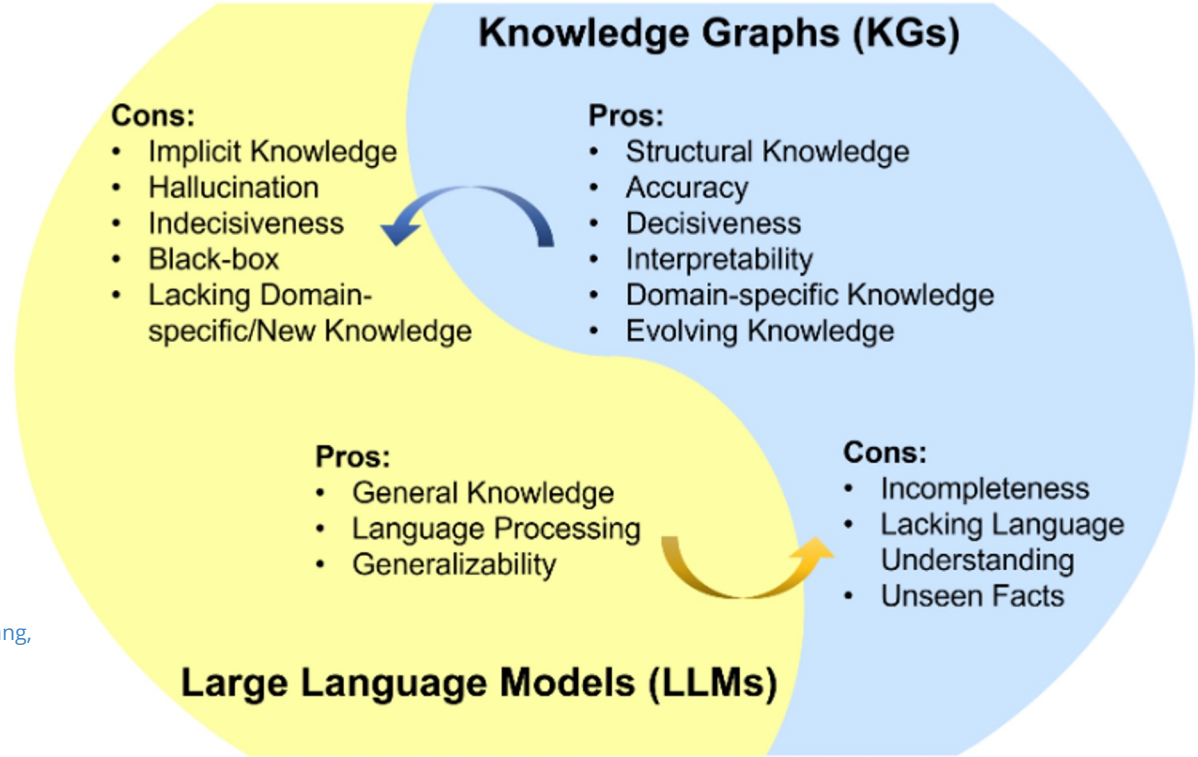
"...AI is the latest technologic advance in medicine, with research studies being conducted across most medical specialties...." >>

Section: Introduction

AI in the Intensive Care Unit: Up-to-Date Review
Diep N. Nguyen[1], Brandon Ngo[2], Eric vanSonnenberg[3] 2020

LLMs + Knowledge graphs

- Gemini (Google) (KG-enhanced LLM)
- Semantic Scholar (LLM-enhanced KG)
- IRIS.AI (for pay)



Source: Pan, Shirui, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu. 2023. 'Unifying Large Language Models and Knowledge Graphs: A Roadmap'. arXiv. https://doi.org/10.48550/arXiv.2306.08302.

Finding AI-based services and tools

- Identify relevant tools
- Provide information about reliable service catalogues
- Maintain internal lists or catalogues of useful and reliable services
- Test before making a decision to use and/or purchase; try to negotiate a longer trial period and involve more people in testing (and experts, if available)
- Ensure access to research tools (check terms and conditions, pricing if applicable, etc.); pay special attention to data security, privacy and copyright

Challenges:

- Disciplinary knowledge is needed
- Difficult to assess paywalled tools based on a demo
- Difficult to assess the sustainability of tools

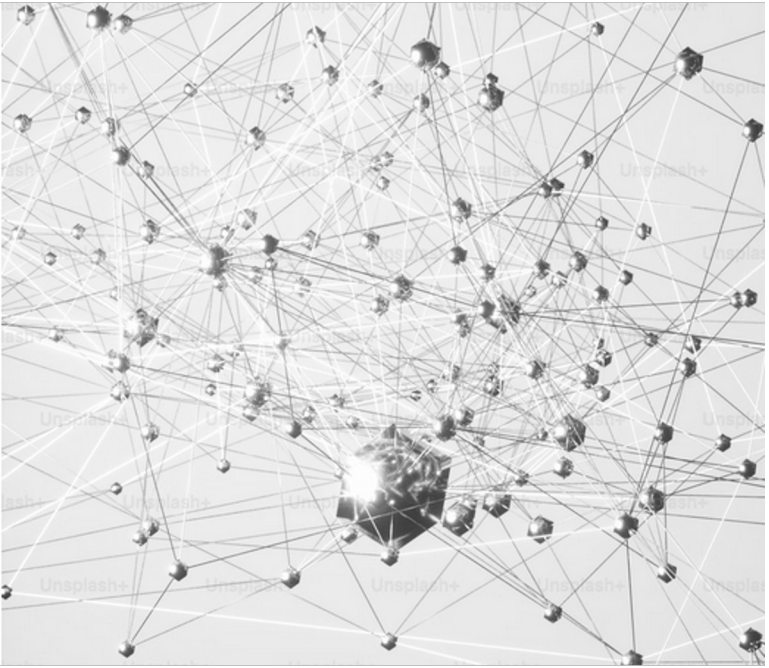
Chatbots and academic tools

- Open source chatbots: https://research.aimultiple.com/open-source-chatbot/
- Elicit https://elicit.com/ (search engine)
- Consensus https://consensus.app/ (search engine)
- Assistant by scite https://scite.ai/assistant (search engine)
- Inciteful https://inciteful.xyz/ - built-in with Zotero
- R Discovery https://discovery.researcher.life
- PaperPal https://paperpal.com/ (writing assistance)
- Jenni https://jenni.ai/ (writing assistance)
- SciSummary https://scisummary.com/ (text summarize)
- Quillbot https://quillbot.com/ (paraphrasing, grammar check, plagiarism check)

Will AI tools make librarians' support in information discovery superfluous?

- Remember "I don't need a library, I have Google."?
- Remember the hype around Web 2.0?
- Poor searching skills among researchers
- Poor knowledge of what to find in particular databases and what searching engines do (and what they don't do)
- Researchers aren't always able to articulate research questions properly
- Limited information literacy (not only among junior researchers)

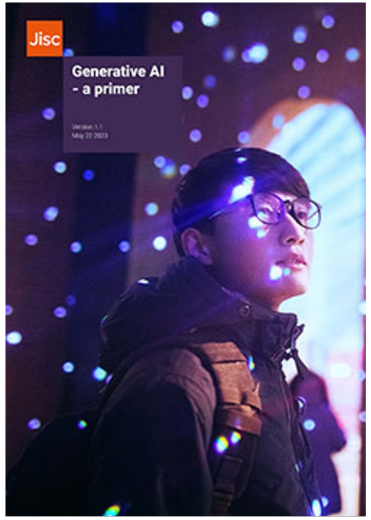
How can research libraries contribute to the development of AI?



- Enabling access to high-quality content for LLM training and knowledge graphs (through repositories and institutional publishing platforms: open access to metadata, publications and data)
- Developing the culture of sharing research outputs
- Developing interoperable infrastructure
- Improving machine readability through metadata and data curation
- Encouraging researchers to use open-source AI tools (to improve their sustainability)

Recommendations

- Living guidelines on the responsible use of generative AI in research (ERA Forum Stakeholders' document), 2024
- Chatbots, ChatGPT, and Scholarly Manuscripts, WAME Recommendations on ChatGPT and Chatbots in Relation to Scholarly Publications
- UNESCO. 2023. 'UNESCO's Recommendation on the Ethics of Artificial Intelligence: Key Facts'. https://unesdoc.unesco.org/ark:/48223/pf0000385082.
- 'WHO Calls for Safe and Ethical AI for Health'. n.d. Accessed 28 November 2023. https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health.
- Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. Zenodo. https://doi.org/10.5281/zenodo.3240529
- Responsible AI licences (open RAIL): https://www.licenses.ai/ai-licenses
- Sabzalieva, Emma, and Arianna Valentini. 2023. 'ChatGPT and Artificial Intelligence in Higher Education: Quick Start Guide - UNESCO Digital Library'. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385146.
- And more from UNESCO: https://www.unesco.org/en/digital-education/artificial-intelligence
- 'Generative AI : A Primer'. 2023. JISC. https://beta.jisc.ac.uk/reports/generative-ai-a-primer.



Questions?

milica.sevkusic@eifl.net

X: @lessormore4