From Policy to Practice: AI Literacy, Quality, and Governance in Academic Libraries

Presented by Dr. Ashleigh Faith, MLIS, PhD

Director of AI and Semantic Innovation at EBSCO

afaith@ebsco.com or
ai@ebsco.com

Follow us on LinkedIn and our EBSCO AI landing page

Generative AI: What are the challenges and opportunities in the research space (and where can/are librarians leading the way)

EBSCO is working with librarians on Standards Committees, IRBs and hosting a new AI for Libraries Discussion Group

Challenges

- Misinformation, hallucinations, inaccurate, incorrect, and inconsistent AI generation
- Information, data, and AI literacy to faculty
- Specificity of AI generation and domain expertise
- Maintaining the rigor of research and librarianship while balancing the efficiencies AI offers
- Responsibility in ethics, protecting against biases, copyright infringement, plagiarism, environmental impact, unbalanced training sets, costs; maintaining privacy
- Rapid advancements in AI; slow advancements in AI standards and regulation

Opportunities

- Grounding AI in the library and academic publishing sources of truth
- Recommendations for content, authors, research topics, subject heading indexing, collection development and deaccessioning
- Insights into content and search; hypothesis tuning and literature review surveys; research findings, topics, and authors and institutions; research habits, library habits, borrowing habits
- Semantic search through honoring user intent and existing query expansion
- Task automation, reference desk assistant chat, copy cataloging, and matching library collection materials to curricula and syllabi
- Translation, breaking down barriers to entry, and improved OCR and digitization of content

Librarian-influenced AI areas :

Open Access is Trending Down; AI Crawling is Trending Up

Trends

- Major research output (23% of all scientific research) from counties without read and publish, plus decreases in grant funding
  - Historically, universities pay for 'read' subscription models for open access publication on behalf of their researchers. Due to this, it is more difficult to negotiate transformative agreements between librarians and publishers at certain universities for their researchers to publish in open access journals.
- AI is increasing usage
  - Library Journal, and many at Charleston Conference 2024 reported AI crawls were so prevalent they were taking down their library websites. Universities are seeing more usage of their catalogs from Google than library search. In some ways, this helps libraries get used if you can get the data up and managed.

Impact for Libraries

- This is a trend toward publishing outside of Open Access, however pre-prints are on the rise for early feedback on publications before submissions. Librarians that oversee digital humanities and catalogs are doing more metadata creation than ever before.
- AI is increasing the user/patron base of library resources, but more effort is needed to establish best practice to manage the influx of crawls.

Trust is Critical, But How to Determine the Source of Truth?

Trends

- Verification Tools

  More emphasis is on verifying the output of the AI since many of the AI models are black boxes. These include MITs SymGen, frameworks such as that outlined by Imperial College and Revenue and Customs UK, or Google Fact Check Explorer.

- Subject Matter Expert Verification

  In some countries, verifying your expertise before weighting in on a topic is required. More and more AI assessments are done via survey tools like Mechanical Turk or Toloka which also require verification of expertise before AI outputs can be assessed.

Impact for Libraries

- High quality metadata and information is needed for fact verification. Librarians understand how to verify information better than anyone, due to information literacy training.

- Librarians have subject matter expertise. We also know how to work with our researchers to verify what AI is doing. This creates a close partnership between researchers, departments, and librarians.

AI Attribution, Copyright, and Terminology

Trends

- AI watermarks, citations, and disclaimers

  More effort is going into how AI should be cited, like the MLA statements on citing and using AI, and when its best to use AI in the research process, in addition to understanding how the AI uses content at a high level.

- "You will know a word by the company it keeps"

  Query expansion has been done through linked data knowledge graphs for a long time, but now AI with these knowledge graphs helps align query intent with the search logic for "deeper" research, as the Bengaluru School of Computer Science and Engineering, and many others, have found.

Impact for Libraries

- AI literacy for when, how, and to what end AI can be used responsibly in AI is a critical skill and role for librarians. Most researchers don't know they shouldn't use content under copyright in AI tools. Helping researchers understand how they can cite/disclose their AI use, and how they can determine if others have used it, is where librarians can help educate

- The taxonomy of AI is nebulous and could do with some librarian cataloger and subject indexer TLC! Librarians can help with defining the logic and terminology for this for their users, and any AI efforts their researchers are engaged in.

Library data, specifically linked data, special collections, and archival data, are in high-demand.

Trends

- Small or fine-tuned models over LLMs

  A study done by NVIDIA and the Georga Institute of Technology has shown that small language models outperform and are more accurate than LLMs when trained on and by specific organizations and their own data. They concluded this would be more economic, and therefore one can conclude more environmentally beneficial, for AI development. Paper.

- World models instead of word models

  MIT, Cornell, and the Basis Research Institute just released the Benchmarking World-Model Learning paper. It unveils WorldTest, a benchmark for assessing: does an AI understand how the world around it works and adapt as it changes? Results, humans still understand the world better than foundational models like OpenAI or Claude. Paper.

Impact for Libraries

- Individuals and organizations taking control of their own data and AI needs increases the need for high quality, trustworthy information, specifically with subject tags.

- Including different cultural, voices, and people groups, preservation of those now and in the past, is something the library has been doing for generations. This is the data needed for world models and helps librarians support more ethical and responsible AI practices.

EBSCO’s AI What we have and where we plan to go:

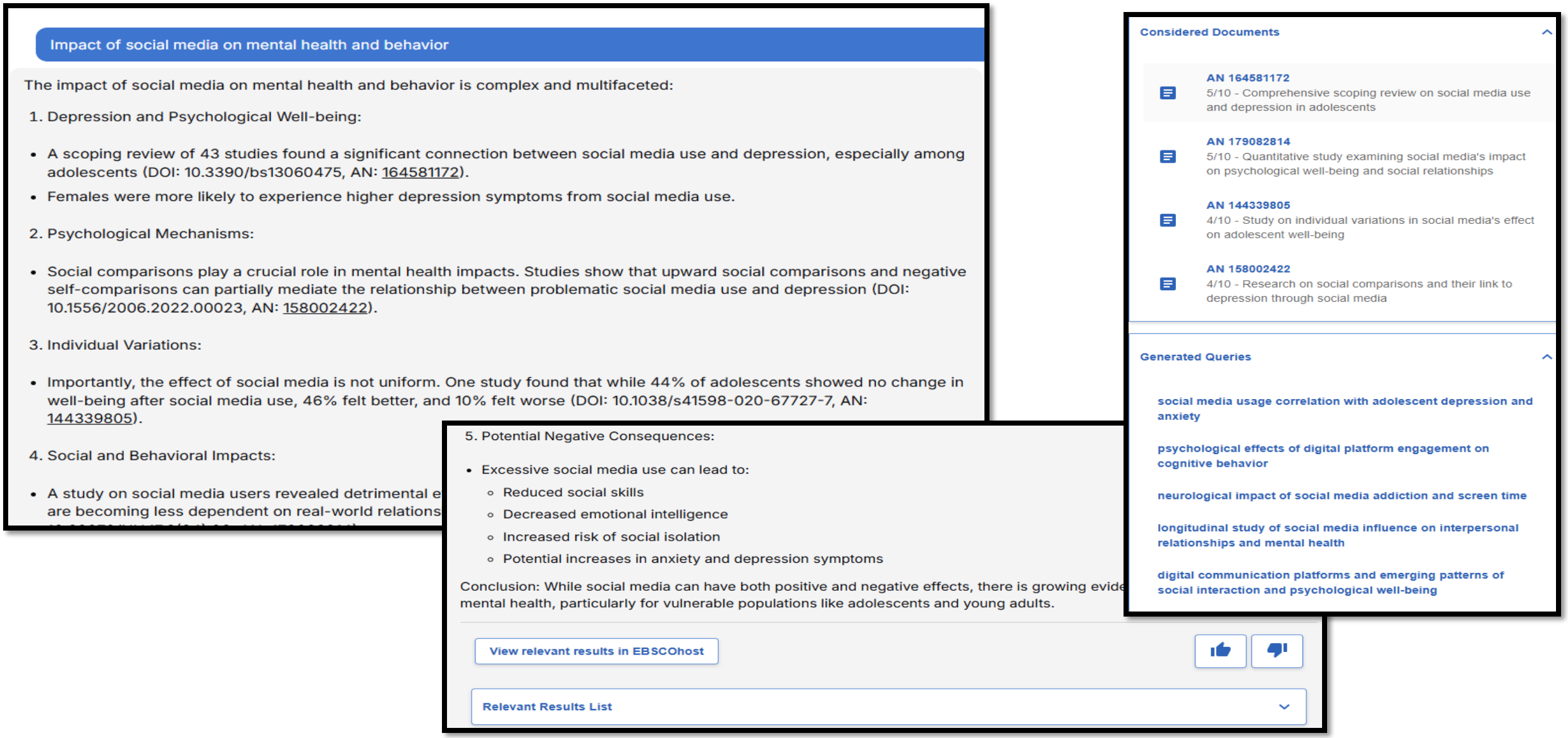
Focused on Responsible Research through Responsible AI

	2024 Betas (complete!)	2025 and Beyond
 Observe and be inspired to ask a question	**Natural Language search**	Search result synthesis; **Special Collections Discovery**
 Create a hypothesis		**AI suggested search recommendations; Recommendations 2.0**
 Test the hypothesis with experimentation and evidence gathering	**Dyna AI Summaries; AI Insights; Natural Language Search**	**AI Assistance w/ Literature Review, Recommendations, Citation Analysis/Bibliometric Pages**
 *Analyze findings through scholarly methods	*We are not currently investigating any AI that would affect the assessment and critical thinking pieces of the research process. Our goal is to retain the use of critical thinking and information (and AI) literacy skills during the researchers journey though EBSCO products.	
 Document findings compared to hypothesis		AI Assistance w/ Literature Review Chat

Plus, Responsible AI internal training, Open and freely accessible AI Literacy Course and AI Transparency documentation.

Conversational Search

AI-powered conversational discovery that integrates natural language processing and machine learning to deliver contextual, responsive search experiences

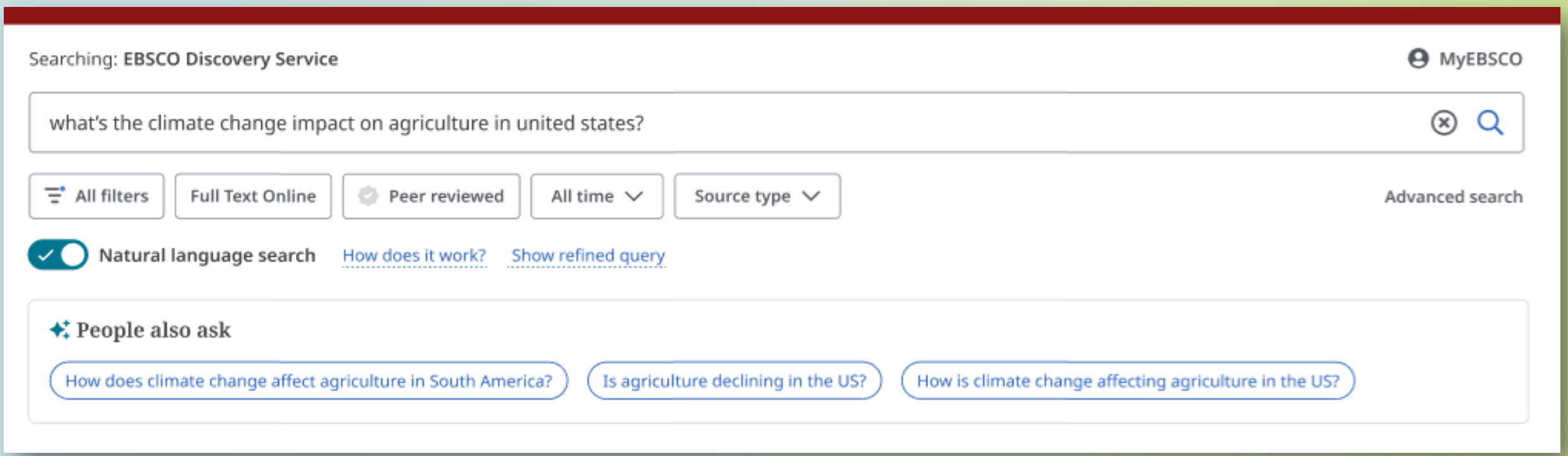


AI Product Features So Far

Available now!

GenAI Search Recommendations

Search Recommendations will help researchers **iteratively drill down into relevant concepts** within the context of the user's original search.



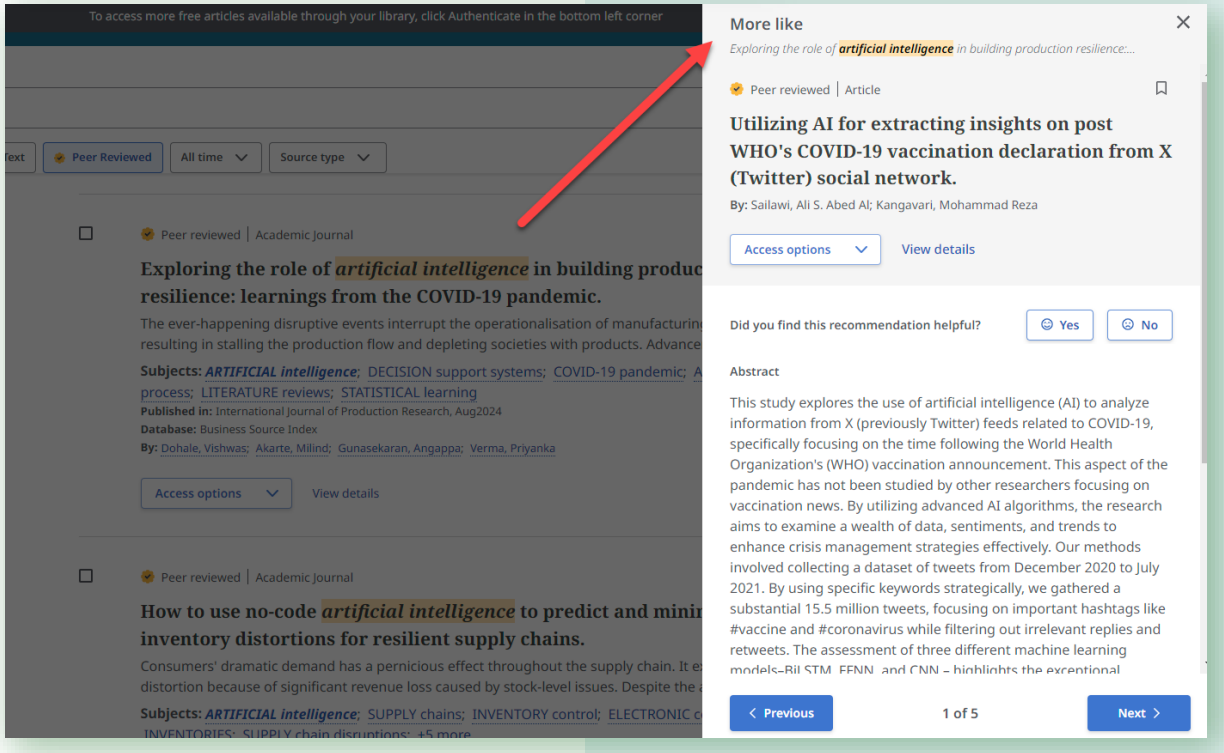
In Beta

Recommendations 2.0

Recommendations 2.0
An updated recommendation engine for content that testing has shown increases relevancy almost 3X as much as the 1.0 version.

Users who click “More Like This” will see a list of recommendations powered by an advanced semantic search & citation network algorithm, powered by the GSN.

Quote from users: “*It can sometimes be really hard to find books for your paper so if I can find one and then potentially find two more or three more, that’s extremely helpful*”



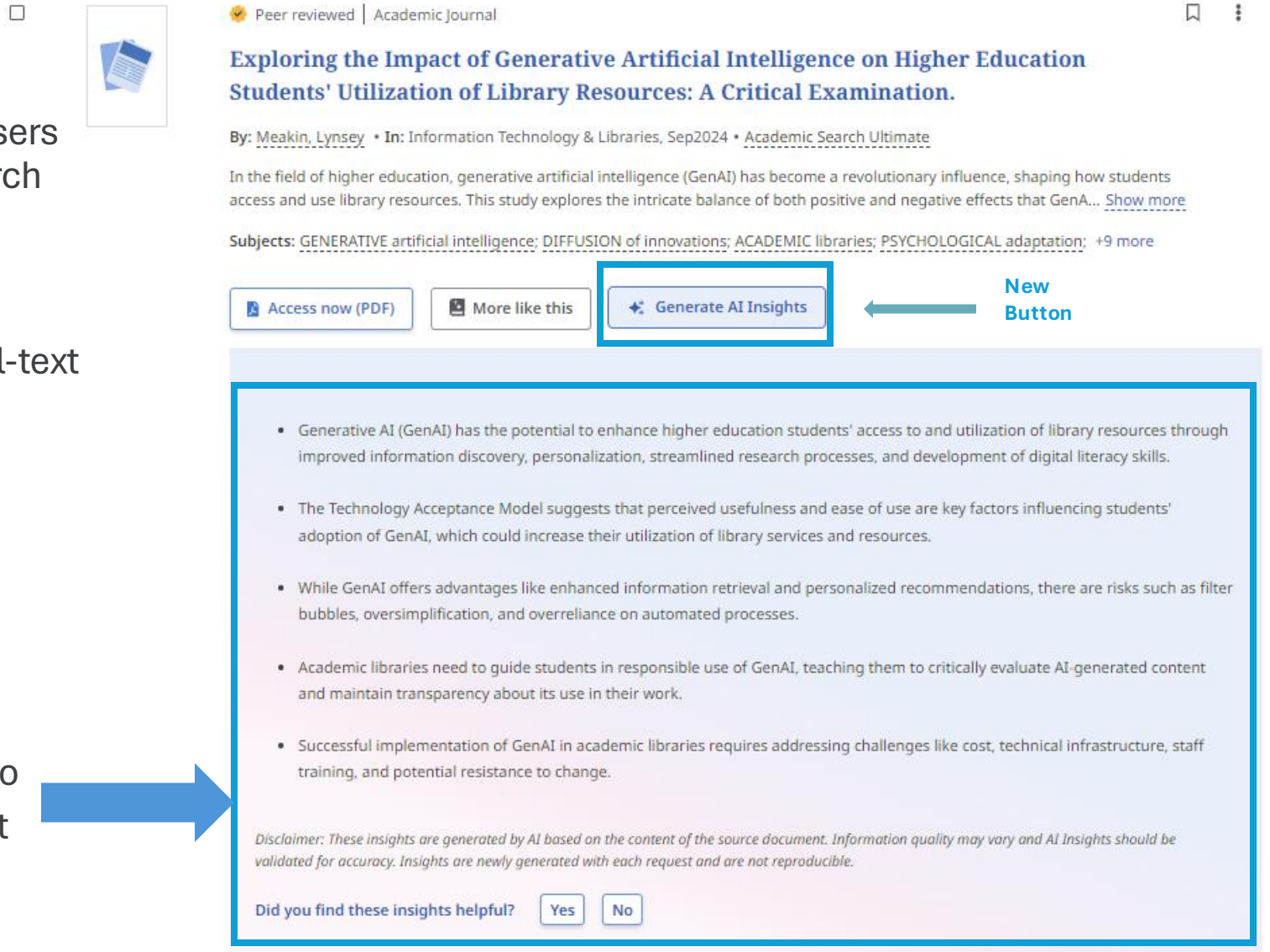
Available now!

Generate *AI Insights* for FT Articles

EBSCO *AI Insights* is a new tool that helps users evaluate relevancy of an article to their research topic

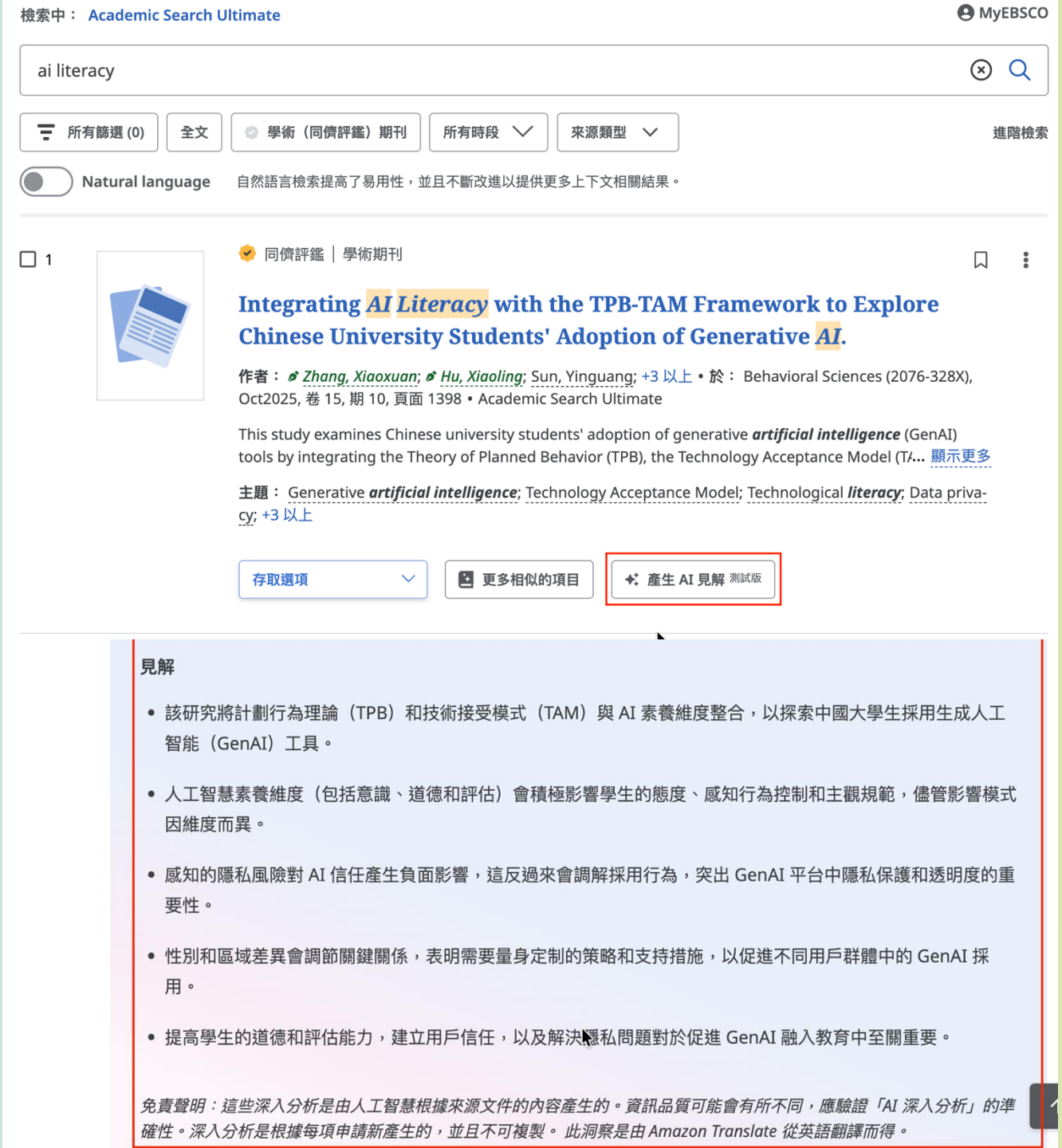
- Leverages Generative AI
- New button on the result list record for full-text records
- Insights are on-demand summaries of 2-5 article key points

Generative AI Insights summaries are clearly marked with a disclaimer encouraging users to validate outputs against the source document



Available now!

AI Insights can transform after switching the interface language.



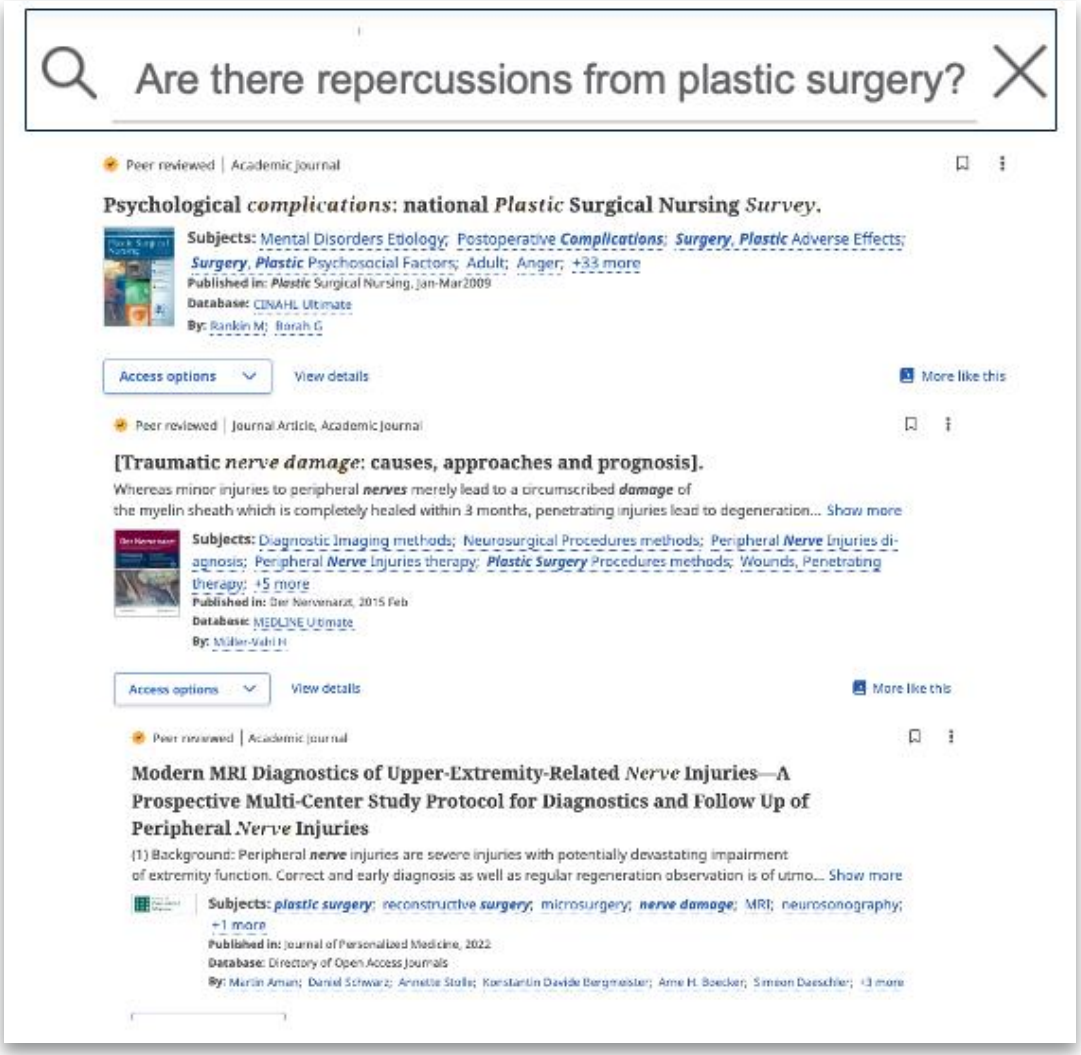
Available now!

Natural Language Search

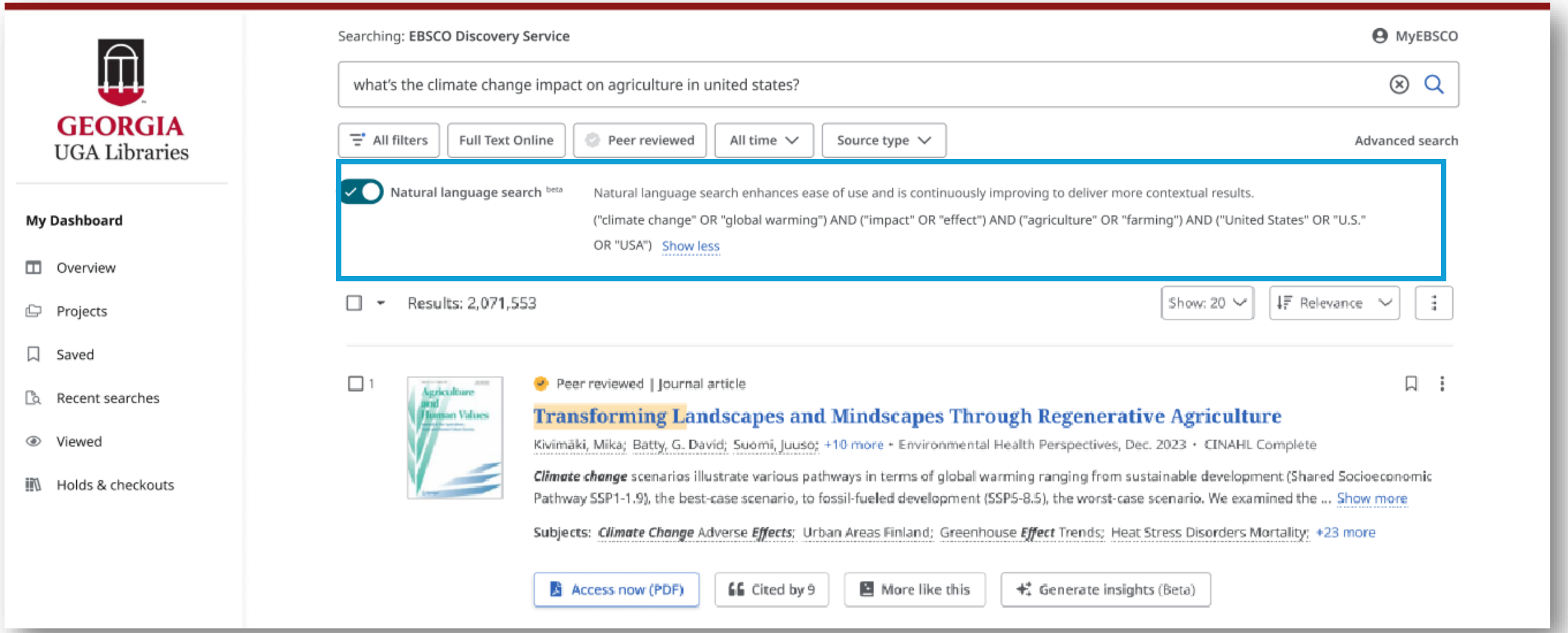
Our upcoming natural language query allows for searching in the researcher's regularly spoken language, including questions.

The underlying search technology leverages our relevance ranking, thesauri mapping via the USI and authoritative metadata

Natural language search is intended to help increase equity in research allowing anyone, regardless of their background, experience in research, or what they speak, can find information through our search.

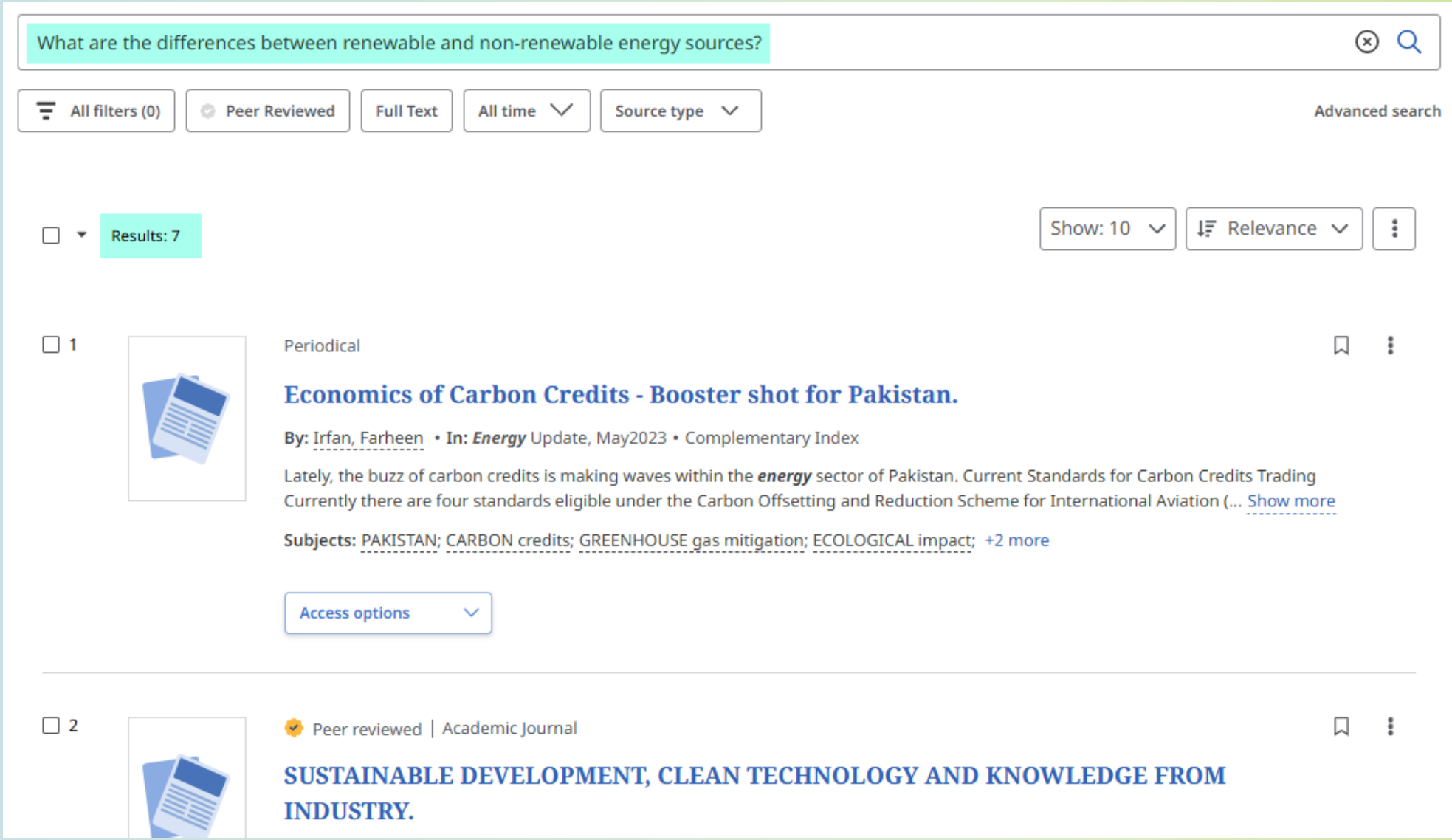


Transparency equivalent Boolean query



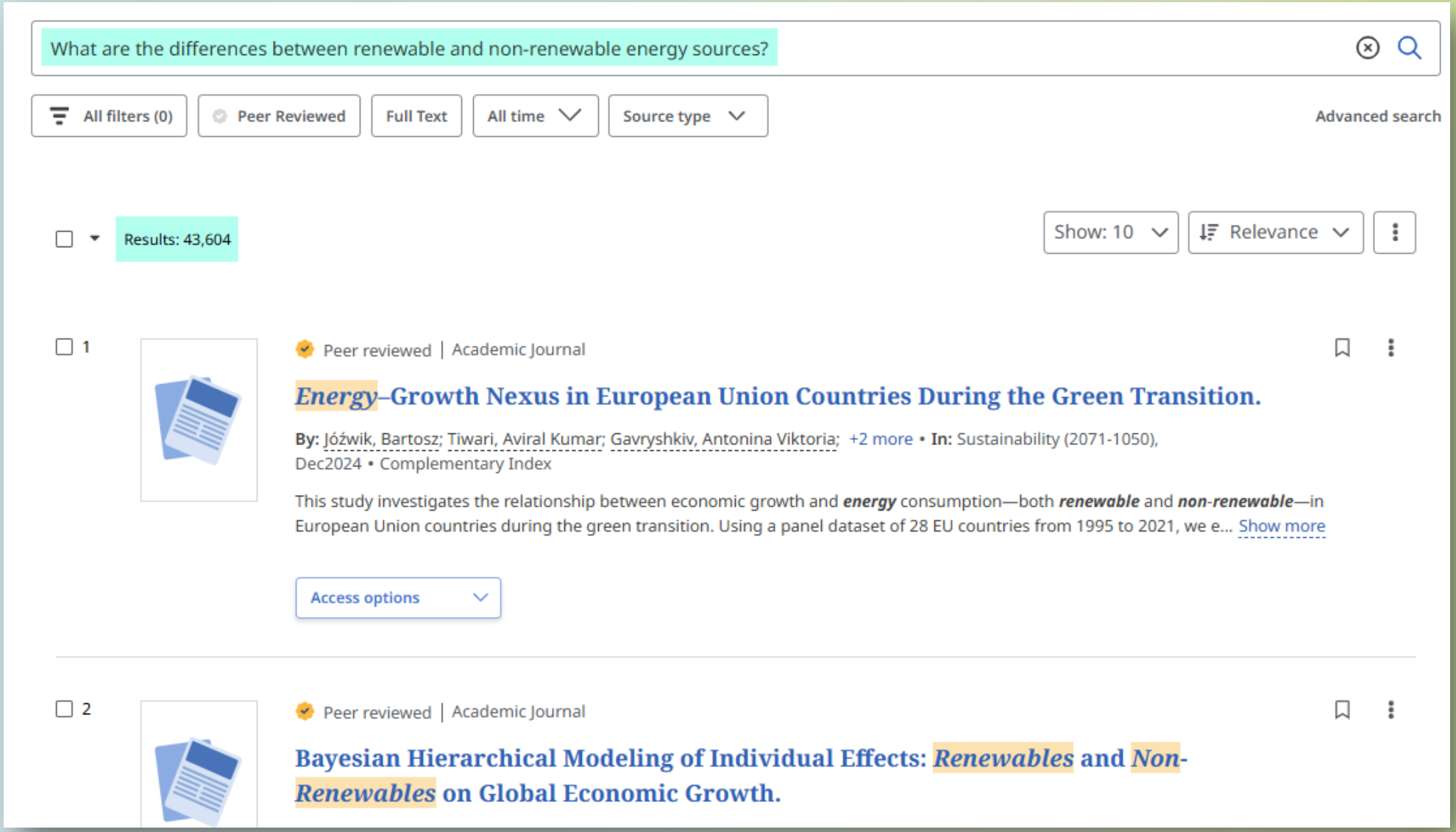
Traditional search

What are the differences between renewable and non-renewable energy sources?



Natural language search

What are the differences between renewable and non-renewable energy sources?



Traditional search

How has technology evolved over the past decade?

How has technology evolved over the past decade?

All filters (0) | Peer Reviewed | Full Text | All time | Source type | Advanced search

Results: 164 | Show: 10 | Relevance

1

Peer reviewed | Academic Journal

Position paper on *how technology* for human motion analysis and relevant clinical applications have *evolved over* the *past decades*: Striking a balance between accuracy and convenience.

By: Bonato, Paolo; Feipel, Véronique; Corniani, Giulia; +2 more • **In:** Gait & Posture, Sep2024 • Academic Search Ultimate

Over the ***past decades***, tremendous technological advances have emerged in human motion analysis (HMA). ***How*** has ***technology*** for analysing human motion ***evolved over*** the ***past decades***, and what clinical applications has it enabled? The literature on H... Show more

Subjects: MOTION analysis; BALANCE disorders; DEEP learning; GAIT disorders; +1 more

Access options | 2 citations in Scopus®

2

Peer reviewed | Academic Journal

Review article: Prehospital telehealth for emergency care: A scoping review.

Natural language search

How has technology evolved over the past decade?

How has technology evolved over the past decade?

All filters (0) | Peer Reviewed | Full Text | All time | Source type | Advanced search

Results: 2,101,153 | Show: 10 | Relevance

1

Peer reviewed | Academic Journal

Revolutionizing Healthcare: Convergence of IoT and Open-Source ERP Systems in Health Information Management.

By: Alihamidi, Imam; Deroussi, Anass; Addaim, Adnane; +1 more • **In:** International Journal of Online & Biomedical ***Engineering***, 2024 • Complementary Index

Over the ***last decade***, health information systems (HIS) have undergone significant ***changes***, particularly in embracing flexible frameworks for ongoing ***development***. This ***evolution*** underscores the necessity for information ***technology*** (IT) infrastruc... Show more

Subjects: All Other Information Services; MANAGEMENT information systems; INFORMATION resources management; INFORMATION ***technology***; +2 more

Access options

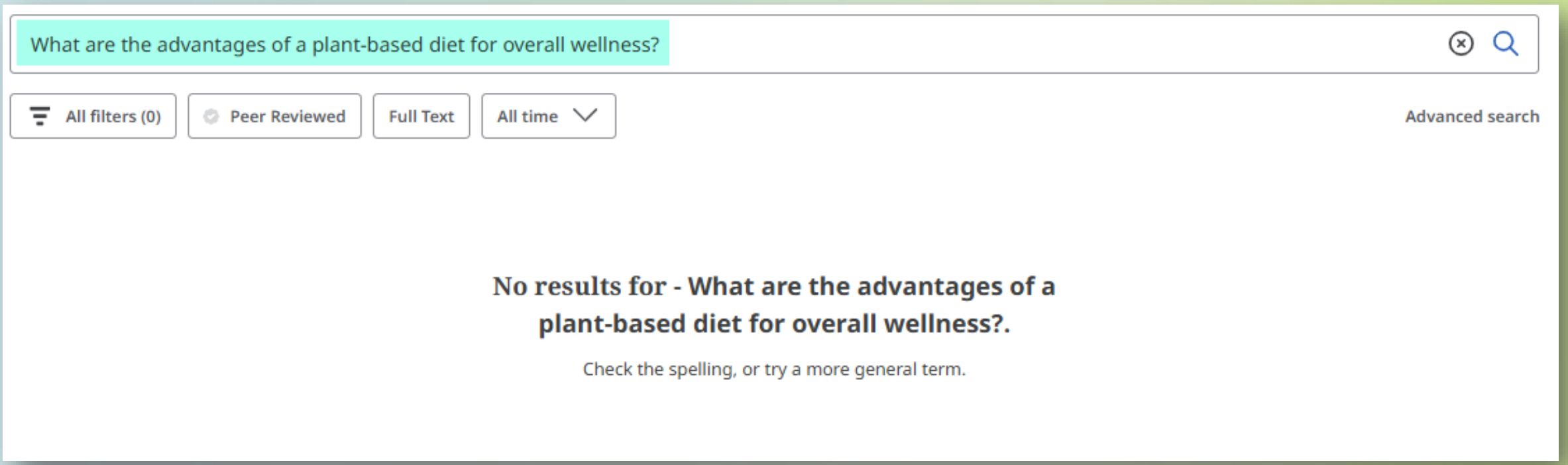
2

Peer reviewed | Academic Journal

Shaping the techno-social landscape of corrections: How values, *technology*, and culture influence the design of correctional service delivery applications

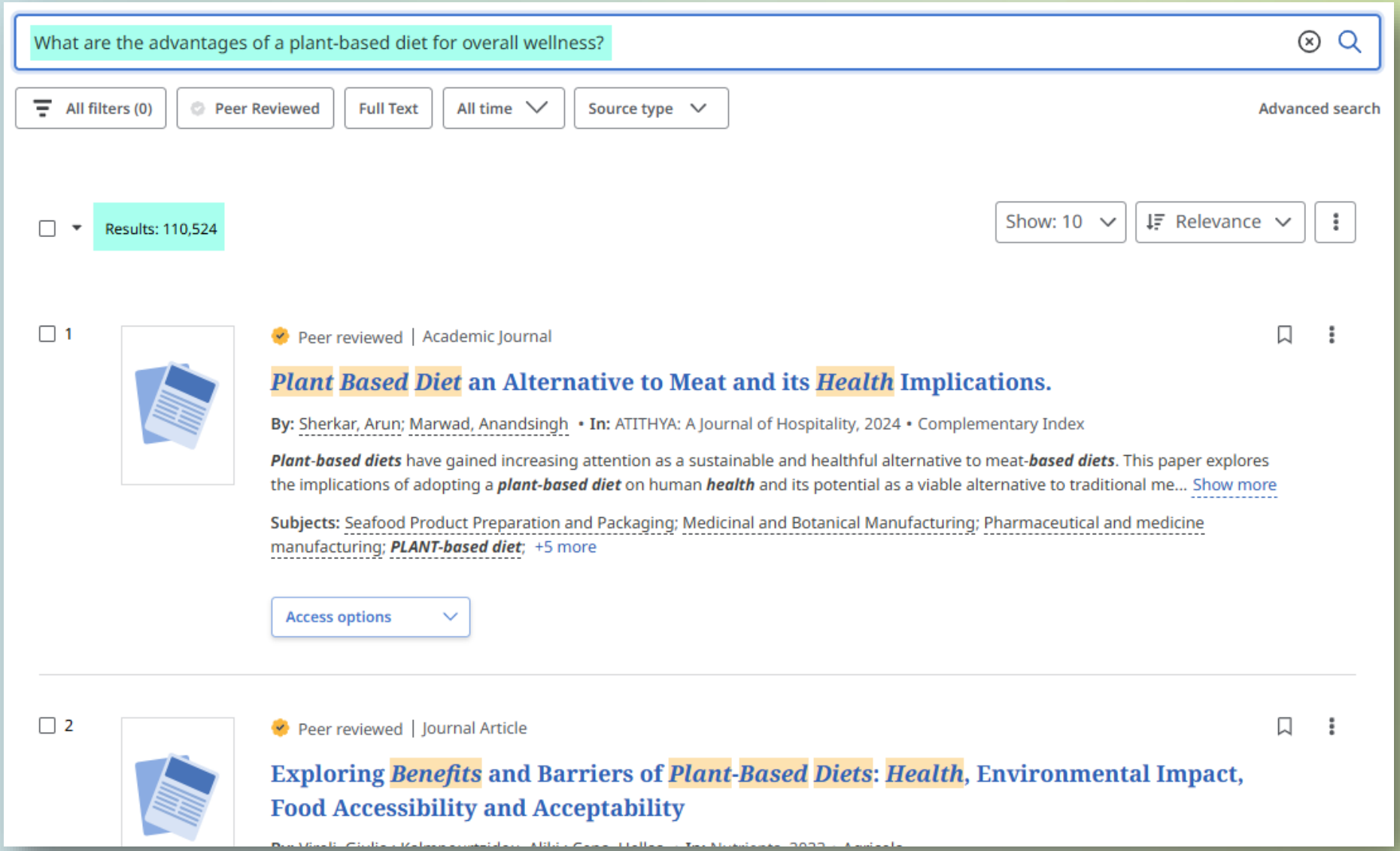
Traditional search

What are the advantages of a plant-based diet for overall wellness?



Natural language search

What are the advantages of a plant-based diet for overall wellness?



Reading Materials

Where can you learn more information about AI and knowledge graphs?

- https://arxiv.org/abs/2311.07509
- https://arxiv.org/abs/1806.01261
- https://arxiv.org/abs/2310.10553
- https://github.com/microsoft/graphrag
- https://neo4j.com/blog/graphrag-manifesto/
- https://neo4j.com/labs/genai-ecosystem/langchain/?gad_source=1&gclid=CjwKCAjwmaO4BhAhEiwA5p4YLziMQQohmxiiDjVqNYlSDfFursX39k32F6cCpUcYi3uZ_Zfr2HWAyRoCtV0QAvD_BwE
- https://paperswithcode.com/paper/head-to-tail-how-knowledgeable-are-large
- https://aclanthology.org/2022.tacl-1.11/
- https://arxiv.org/abs/2306.08302
- https://online.stanford.edu/courses/xcs224w-machine-learning-graphs
- https://crfm.stanford.edu/fmti/May-2024/index.html
- https://open.hpi.de/courses/knowledgegraphs2020
- https://learn.deeplearning.ai/courses/knowledge-graphs-rag/lesson/4/preparing-text-for-rag
- https://www.techtarget.com/searchcontentmanagement/tip/How-a-content-tagging-taxonomy-improves-enterprise-search
- https://medium.com/@joehoeller/understanding-graph-types-and-ontological-driven-data-structures-185eb176c3c3

- https://enterprise-knowledge.com/generative-ai-for-taxonomy-creation/
- https://cltc.berkeley.edu/publication/a-taxonomy-of-trustworthiness-for-artificial-intelligence/
- https://www.linkedin.com/learning-login/share?forceAccount=false&redirect=https%3A%2F%2Fwww.linkedin.com%2Flearning%2Fknowledge-graphs-for-generative-ai-use-cases%3Ftrk%3Dshare_ent_url%26shareId%3DJjgNbhXrQWiBI8cwRtibZw%253D%253D
- https://www.youtube.com/watch?v=pjxzAczeuNo
- https://youtu.be/ZOO24NQHZfM
- https://youtu.be/FJ3ZqtnJ_go
- https://youtu.be/urTtHSXaab0
- https://youtu.be/B3F3V569xKQ
- https://www.hedden-information.com/wp-content/uploads/2022/03/Getting_Started_with_Taxonomies.pdf
- https://www.databricks.com/blog/LLM-auto-eval-best-practices-RAG
- https://arxiv.org/abs/2304.10145
- https://arxiv.org/pdf/2404.11794
- https://www.linkedin.com/pulse/spellbinding-tech-tale-rag-kg-ashish-singh-zuosc/
- https://ieeexplore.ieee.org/abstract/document/10417790\

Thank you!