

But where are we marching to?

Of trust and provenance: knowing where we've been and deciding where we're going.

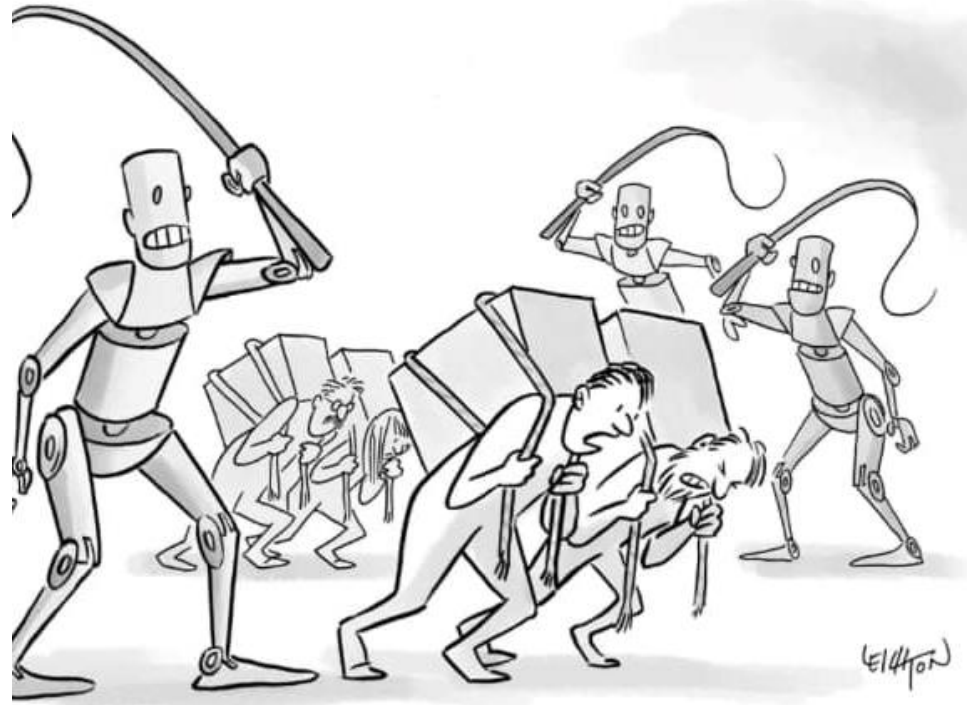
Anita de Waard

VP Research Collaborations

a.dewaard@elsevier.com



Fiesole
Retreat 23



“To think this all began with letting autocomplete finish our sentences.”

My Abstract! *What do you think? Any issues?*

From fake news to papermills to plagiarizing chatbots, the knowledge objects that scholars, libraries and publishers, create, share and curate are subject to doubt. So how can we restore confidence in research? AI is here to stay, and will increasingly gain ground as a partner in the research process. But we need to make sure that what we create and curate together in this cyborgian knowledge environment serves humanity, and the future we hope to create. This talk will raise a few questions and share a few examples on how<.>we can use AI to combat the challenges of doubt and misinformation in research. I'll discuss how AI can be used to verify sources, detect plagiarism and generate trustworthy research outputs. Additionally, we will explore the potential of AI to assist with data analysis, identify patterns and generate new insights. However, as we embrace AI in research, we must also be mindful of the ethical implications and potential biases that could arise. Ultimately, we need to work collaboratively to ensure that AI is used in a way that aligns with our values and contributes to a sustainable and equitable future for all.

My and ChatGPTs Abstract!

From fake news to papermills to plagiarizing chatbots, the knowledge objects that scholars, libraries and publishers, create, share and curate are subject to doubt. So how can we restore confidence in research? AI is here to stay, and will increasingly gain ground as a partner in the research process. But we need to make sure that what we create and curate together in this cyborgian knowledge environment serves humanity, and the future we hope to create. This talk will raise a few questions and share a few examples on how<.>we can use AI to combat the challenges of doubt and misinformation in research. I'll discuss how AI can be used to verify sources, detect plagiarism and generate trustworthy research outputs. Additionally, we will explore the potential of AI to assist with data analysis, identify patterns and generate new insights. However, as we embrace AI in research, we must also be mindful of the ethical implications and potential biases that could arise. Ultimately, we need to work collaboratively to ensure that AI is used in a way that aligns with our values and contributes to a sustainable and equitable future for all. </.>

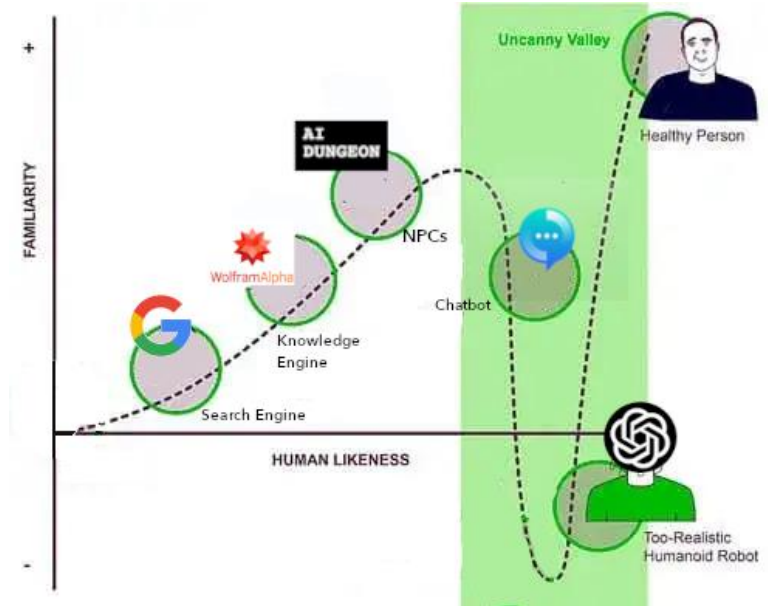


Is that a problem for you?




Uncanny valley for language!

Language models can generate language answers that are either perfectly human and helpful or wildly inhuman or incorrect, without clear reasons as to which type of answer one will receive, leading to discomfort and distrust of the technology.

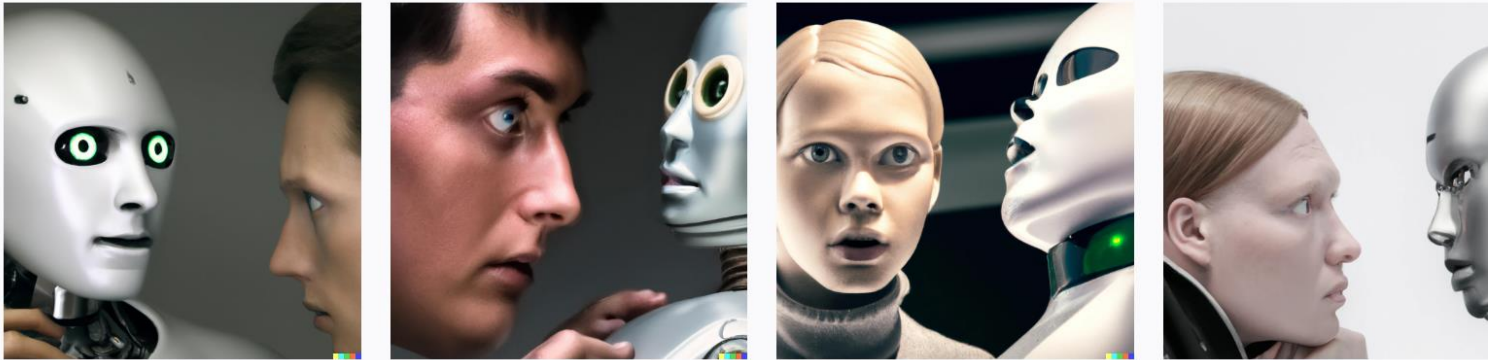


Actually, I made that picture... Or did I?

 DALL-E History Collections

Edit the detailed description Surprise me Upload

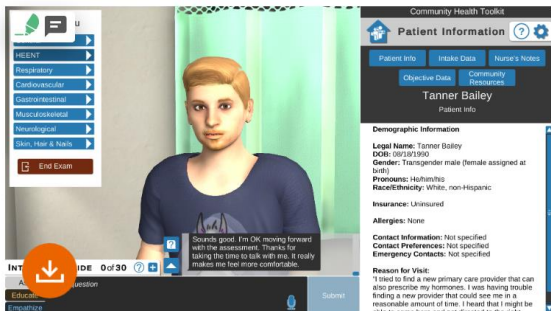
Uncanny valley: a photograph of a cyborg that looks very realistic. A person is looking at it, very afraid. Generate



The image displays four generated photographs illustrating the uncanny valley concept. Each image shows a human face in profile or three-quarter view, looking at a highly realistic-looking robot. The robots have human-like features but with subtle, unsettling differences, such as glowing eyes or unnatural skin textures. The human faces exhibit expressions of concern, fear, or intense focus.

- Copyright? DOI? Ownership?
- How do we archive, index, sell, share... publish this?

AI can do great things for scholarly communications:



Download

Impact of a Virtual Patient Simulation on Nursing Students' Attitudes of Transgender Care

Study findings support experiential learning as an effective tool in teaching cultural competence and caring for transgender patients.

Competition



LaySumm (The 1st Computational Linguistics Lay Summary Challenge Shared Task)

Organized by Abhilasha Ravichander - Current server time: April 27, 2023, 7:32 p.m. UTC

Current

End

First phase

Competition Ends

July 15, 2020, midnight UTC

Never

Learn the Details

Phases

Participate

Results

Overview

Evaluation

Terms and Conditions

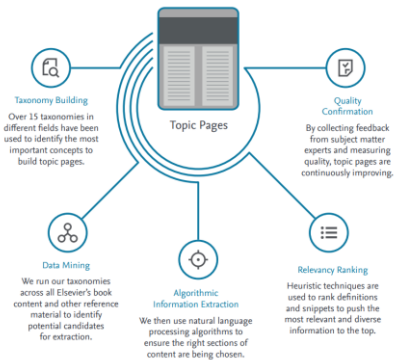
LaySumm - Overview

Lay Summary Generation is a shared task that focuses on generating lay summaries for scientific documents. LaySumm is one of three shared tasks conducted as part of: 1st Workshop on Scholarly Document Processing

AI can help explain scholarly outputs

AI can help protect patients and train nurses

A New Age of Discovery: Using AI to Generate Topic Pages



- https://journals.lww.com/nurseeducatoronline/Citation/2023/05000/Impact_of_a_Virtual_Patient_Simulation_on_Nursing.6.aspx
- https://www.elsevier.com/_data/assets/pdf_file/0006/901635/ACAD_L_SD_INFO_TechSDTopics_WEB.pdf
- <https://competitions.codalab.org/competitions/25516>

But AI can also be biased:



Bias in **data** [Scheuermann et al]:

A particularly interesting example was PUBFIG, which had two gendered annotations: “male” and “attractive woman,” of which there was no associated “female” [87]. The absence of annotations for “female” or “attractive man”, however, highlights the culturally-situated values around gender that can emerge within an annotation schema (c.f., [150]).

Bias in **algorithms** [Amnesty International Report]:

Xenophobic machines: Discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal

October 25, 2021, Index Number: EUR 35/4686/2021

EUROPE AND CENTRAL ASIA **TECHNOLOGY AND HUMAN RIGHTS**

Social security enforcement agencies worldwide are increasingly automating their processes in the hope of detecting fraud. The Netherlands is at the forefront of this development. The Dutch tax authorities adopted an algorithmic decision-making system to create risk profiles of individuals applying for childcare benefits in order to detect inaccurate and potentially fraudulent applications at an early stage. Nationality was one of the risk factors

And AI can be weaponized:

This author profile is generated by Scopus. [Learn more](#)

ChatGPT, null

OpenAI, Inc., San Francisco, United States

[View more](#)

34

Citations by 30 documents

3

Documents

[Set alert](#)

[Edit profile](#)

ORCID

Connecting research and researchers

[ABOUT](#) [FOR RESEARCHERS](#) [MEMBERSHIP](#) [DISCOVER](#)

Search

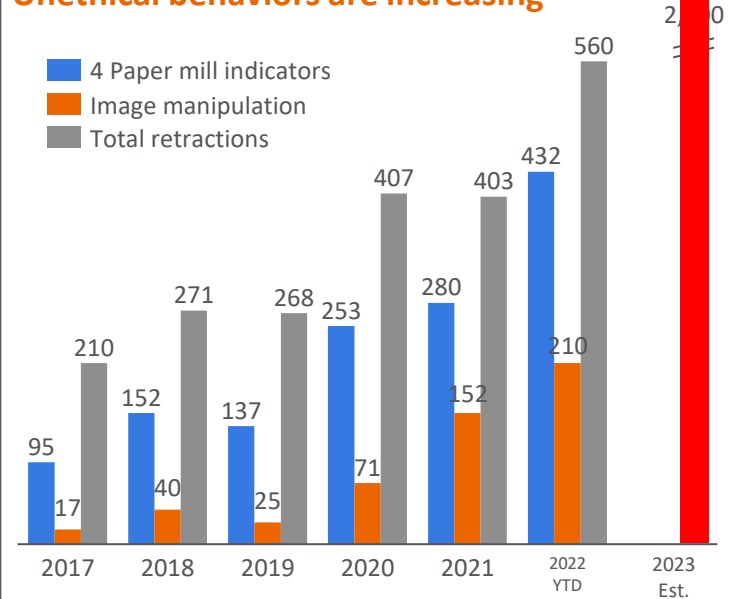
Showing 50 of 557 results.

Items per page: 50 Page 1 of 12

ORCID ID	First Name	Last Name
0009-0006-3494-5979	ChatGPT	ChatGPT
0009-0006-6767-8424	ChatGPT	Online
0009-0003-5955-9546	Chatgpt	Prompts
0009-0006-7009-9386	GPTGo	- Free ChatGPT & Google
0009-0009-8523-5987	GooGPT - Google	Search and ChatGPT

Unethical behaviors are increasing

4 Paper mill indicators
Image manipulation
Total retractions



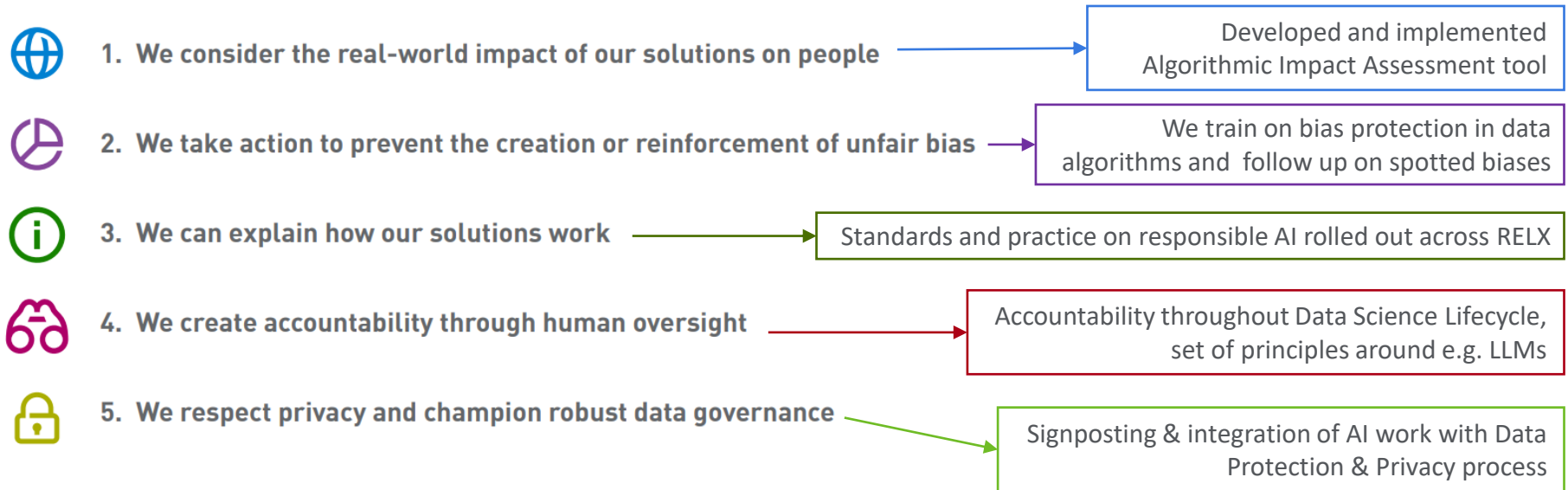
Growth of retractions, 2017-2022 YTD & estimated 2023

- <https://www.the-geyser.com/chatgpt-says-its-not-an-author/> -- <https://www.elsevier.com/about/policies/publishing-ethics/the-use-of-ai-and-ai-assisted-writing-technologies-in-scientific-writing>
- Internal data but covered by Sarah Jenkins at <https://www.stm-assoc.org/events/stm-research-integrity-master-class-2/>

What to do? RELX responsible AI principles:

AS PART OF OUR RESPONSIBLE AI APPROACH...

Here's what that looks like in practice:



What to do?

Collaborative Research!

Overview of the DAGPap22 Shared Task on Detecting Automatically Generated Scientific Papers

Yury Kashnitsky, Drahomira Herrmannova, Anita de Waard, George Tsatsaronis, Catriona Catriona Fennell, Cyril Labbe

"[...] we are far from developing a general detector of generated content. Each new model (say, GPT-4) for which we don't have training data poses a new challenge, and any detector is likely to fail at identifying content generated with such a model..."



- <https://aclanthology.org/2022.sdp-1.26/>
- <https://www.stm-assoc.org/stm-integrity-hub>
- <https://www.niso.org/press-releases/2021/07/stms-peer-review-taxonomy-be-formalized-ansiniso-standard/>



A Standard Terminology for Peer Review

1 Identity transparency:	2 Reviewer interacts with:	3 Review information published:	4 Post publication commenting:																																														
<p>This category describes the extent to which identities of participants are made visible to each other during the review process. Identities not made visible during the process can be made visible in publication on the article page (see table 3). Please note that for this and the other tables the information in bold should be used in communication.</p> <table border="1"><thead><tr><th>TYPE</th><th>DESCRIPTION</th></tr></thead><tbody><tr><td>All identities visible</td><td>Reviewer identity is visible to author; author identity is visible to reviewer; reviewer and author identity is visible to (decision-making) editor</td></tr><tr><td>Single anonymized</td><td>Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer and author identity is visible to (decision-making) editor</td></tr><tr><td>Double anonymized</td><td>Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer and author identity is visible to (decision-making) editor</td></tr><tr><td>Triple anonymized</td><td>Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer & author identity is not made visible to (decision-making) editor</td></tr></tbody></table>	TYPE	DESCRIPTION	All identities visible	Reviewer identity is visible to author; author identity is visible to reviewer; reviewer and author identity is visible to (decision-making) editor	Single anonymized	Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer and author identity is visible to (decision-making) editor	Double anonymized	Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer and author identity is visible to (decision-making) editor	Triple anonymized	Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer & author identity is not made visible to (decision-making) editor	<p>This category relates to direct interaction or exchange of information (e.g. via submission systems or email) during the peer review process. Multiple types of this category can be selected, where applicable. Whatever is communicated about the review process after publication is covered in the category 'Reviewer Information Published'.</p> <table border="1"><thead><tr><th>TYPE</th><th>DESCRIPTION</th></tr></thead><tbody><tr><td>Editor</td><td>Communication between editor and reviewer (traditional model). Also known as 'Independent review'. Identities can be anonymized or visible</td></tr><tr><td>Other Reviewer(s)</td><td>Direct interaction/collaboration (e.g. via submission system or email) between reviewers, or the possibility to receive and/or comment on each other's reports before reviewer makes recommendation to the editor. Identities can be anonymized or visible</td></tr><tr><td>Authors</td><td>Direct interaction/collaboration (e.g. via submission system or email) between author and reviewer before reviewer makes recommendation to the editor. Identities can be anonymized or visible</td></tr></tbody></table>	TYPE	DESCRIPTION	Editor	Communication between editor and reviewer (traditional model). Also known as 'Independent review'. Identities can be anonymized or visible	Other Reviewer(s)	Direct interaction/collaboration (e.g. via submission system or email) between reviewers, or the possibility to receive and/or comment on each other's reports before reviewer makes recommendation to the editor. Identities can be anonymized or visible	Authors	Direct interaction/collaboration (e.g. via submission system or email) between author and reviewer before reviewer makes recommendation to the editor. Identities can be anonymized or visible	<p>This relates to information that is published about the review process on the article page. Select and list the items that are applicable.</p> <table border="1"><thead><tr><th>TYPE</th><th>DESCRIPTION</th></tr></thead><tbody><tr><td>None</td><td>No information about the review process or editorial decision process is published</td></tr><tr><td>Review summaries</td><td>Can be summaries or parts of the minutes, or a summary of the review process</td></tr><tr><td>Reviewer reports</td><td>Full content of the reviewer reports is published.</td></tr><tr><td>Reviewer reports author opt in</td><td>Full content of the reviewer reports is published if the corresponding author opts for this</td></tr><tr><td>Reviewer reports reviewer opt in</td><td>Full content of the reviewer reports is published if the reviewer(s) opt(s) for this.</td></tr><tr><td>Submitted manuscript author opt in</td><td></td></tr><tr><td>Author/Editor communication</td><td>Including editor decision letter and reviewer responses (rebuttals)</td></tr><tr><td>Reviewer identities</td><td></td></tr><tr><td>Reviewer identities reviewer opt in</td><td></td></tr><tr><td>Editor identities</td><td>Identities of the handling editors</td></tr></tbody></table>	TYPE	DESCRIPTION	None	No information about the review process or editorial decision process is published	Review summaries	Can be summaries or parts of the minutes, or a summary of the review process	Reviewer reports	Full content of the reviewer reports is published.	Reviewer reports author opt in	Full content of the reviewer reports is published if the corresponding author opts for this	Reviewer reports reviewer opt in	Full content of the reviewer reports is published if the reviewer(s) opt(s) for this.	Submitted manuscript author opt in		Author/Editor communication	Including editor decision letter and reviewer responses (rebuttals)	Reviewer identities		Reviewer identities reviewer opt in		Editor identities	Identities of the handling editors	<p>Relates to comments on the online published version of the article. Article types such as comment / reply / letter are not considered post-publication commenting as they are stand alone publications. Only use this category when applicable.</p> <table border="1"><thead><tr><th>TYPE</th><th>DESCRIPTION</th></tr></thead><tbody><tr><td>Open</td><td>Commenting open to anybody. Can be anonymous, require signing in and/or registration (e.g. via ORCID)</td></tr><tr><td>On invitation</td><td>Only editor (or publisher) selected and/or invited individuals can comment on the article post-publication</td></tr></tbody></table> <p>Use of terminology</p> <p>As an example, the description of a (traditional) review process by authors (e.g. on the Guide for Authors) would be:</p> <ul style="list-style-type: none">Identity transparency: Single anonymizedReviewer interacts with: EditorReview information published: None <p>Publishers can include links in these descriptions which lead to a page where the terms are explained. In case journals will allow authors to choose between review models, all the options should be listed, e.g.:</p> <ul style="list-style-type: none">Identity transparency: Single anonymized, double anonymizedOnly in the case of post-publication commenting, this should be actively communicated. For example:Identity transparency: All identities visibleReviewer interacts with: Editor, Other Reviewer(s)Review information published: Review reports, reviewer identitiesPost-publication commenting: Open	TYPE	DESCRIPTION	Open	Commenting open to anybody. Can be anonymous, require signing in and/or registration (e.g. via ORCID)	On invitation	Only editor (or publisher) selected and/or invited individuals can comment on the article post-publication
TYPE	DESCRIPTION																																																
All identities visible	Reviewer identity is visible to author; author identity is visible to reviewer; reviewer and author identity is visible to (decision-making) editor																																																
Single anonymized	Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer and author identity is visible to (decision-making) editor																																																
Double anonymized	Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer and author identity is visible to (decision-making) editor																																																
Triple anonymized	Reviewer identity is not made visible to author; author identity is not made visible to reviewer; reviewer & author identity is not made visible to (decision-making) editor																																																
TYPE	DESCRIPTION																																																
Editor	Communication between editor and reviewer (traditional model). Also known as 'Independent review'. Identities can be anonymized or visible																																																
Other Reviewer(s)	Direct interaction/collaboration (e.g. via submission system or email) between reviewers, or the possibility to receive and/or comment on each other's reports before reviewer makes recommendation to the editor. Identities can be anonymized or visible																																																
Authors	Direct interaction/collaboration (e.g. via submission system or email) between author and reviewer before reviewer makes recommendation to the editor. Identities can be anonymized or visible																																																
TYPE	DESCRIPTION																																																
None	No information about the review process or editorial decision process is published																																																
Review summaries	Can be summaries or parts of the minutes, or a summary of the review process																																																
Reviewer reports	Full content of the reviewer reports is published.																																																
Reviewer reports author opt in	Full content of the reviewer reports is published if the corresponding author opts for this																																																
Reviewer reports reviewer opt in	Full content of the reviewer reports is published if the reviewer(s) opt(s) for this.																																																
Submitted manuscript author opt in																																																	
Author/Editor communication	Including editor decision letter and reviewer responses (rebuttals)																																																
Reviewer identities																																																	
Reviewer identities reviewer opt in																																																	
Editor identities	Identities of the handling editors																																																
TYPE	DESCRIPTION																																																
Open	Commenting open to anybody. Can be anonymous, require signing in and/or registration (e.g. via ORCID)																																																
On invitation	Only editor (or publisher) selected and/or invited individuals can comment on the article post-publication																																																

Ecosystem participants, unite!

Introducing the STM Integrity Hub.

The STM Integrity Hub is a robust, direct response to safeguard the integrity of science. Through a combination of shared data and experiences, and by harnessing technological innovation, the STM Integrity Hub offers a holistic approach to detect research-integrity-offending manuscripts. Here, publishers can effectively and efficiently respond to the increasing and alarming volume of materials entering scholarly communications that violate accepted research integrity.

More specifically, the STM Integrity Hub will provide a cloud-based environment for publishers to **check submitted articles for research integrity issues**, consistent with applicable laws and industry best practice and fully respecting the laws and ethics of data privacy and competition/anti-trust laws. In this environment, publishers may **collaborate with other parties of their choosing to develop and operate screening tools** for the benefit of the entire scholarly ecosystem.

Think of it as a **knowledge exchange**: where publishers may share experiences and learnings regarding how best to safeguard research integrity in science; a **think tank** for policy and legal frameworks; and a **living library** of infrastructure and tools.



Let's keep track of who did what to whom and when...

Says **AI Tool X** on **DATE Y**

Be explicit about user of AI!

responses to this virus.
The first steps for such an understanding are the ability to quantify the virus-specific CD4⁺ and CD8⁺ T cells. Such knowledge is of immediate relevance, as it will provide insights into immunity and pathogenesis of SARS-CoV-2 infection, and the same knowledge will assist vaccine design and evaluation of candidate vaccines. Estimations of immunity are also central to epidemiological model calibration of future social distancing pandemic control measures (Kissler et al., 2020). Such projections are dramatically

Based on PAPER P

Says **Editor Z** on **DATE V**

No longer valid

Article
Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals

Alba Grifoni,¹ Daniela Wei,¹ Carolyn Rydzynski Moder,¹ Ramesh S. Jasti,³ Daniel A. Bjoern Peters,^{1,4} Florian K. Thiel,¹ and Florian K. Thiel,¹ Center for Infectious Disease,¹ Department of Medicine, DWR 92037, USA,² Department of Microbiology,³ Department of Microbiology,⁴ These authors contributed equally to this work.
 Lead Contact
 *Correspondence: shane@nih.gov
<https://doi.org/10.1016/j.cell.2020.05.015>

SUMMARY
 Understanding adaptive immunity to SARS-navirus disease 2019 (COVID-19) pathogeni class I and II predicted peptide "megapo were identified in ~70% and 100% of COVID to spike, the main target of most vaccine eff SARS-CoV-2 IgG and IgA titers. The M, spike response, with additional responses comm CD8⁺ T cells, spike and M were recognized, detected SARS-CoV-2-reactive CD4⁺ T cell reactive T cell recognition between circulat

different depending on whether SARS-CoV-2 infection creates substantial immunity, and whether any cross-reactive immunity exists between SARS-CoV-2 and circulating seasonal "common cold" human coronaviruses. Definition and assessment of human antigen-specific SARS-CoV2 T cell responses are best made with direct *ex vivo* T cell assays using broad-based epitope pools and assays capable of detecting T cells of any cytokine polarization. Herein, we have completed such an assessment with blood samples from COVID-19 patients. There is also great uncertainty about whether adaptive immune responses to SARS-CoV-2 are protective or pathogenic, or whether both scenarios can occur depending on timing, composition, or magnitude of the adaptive immune response. Hypotheses range the full gamut (Peeples, 2020), based on available clinical data from severe acute respiratory disease syndrome (SARS) or Middle East respiratory syndrome (MERS) (Aishukairi et al., 2018; Wong et al., 2004; Zhao et al., 2017) or animal model data with SARS in mice (Zhao et al., 2009, 2010, 2016), SARS in non-human primates (NHPs) (Liu et al., 2019; Takanano et al., 2008) or feline infectious peritonitis virus (FIPV) in cats

55425663600
 This is author Alba Grifoni

hggraphid:2ab6cd87-e543-3229-85ff-c862a90f415c (SARS-CoV2)
 hggraphid:88f9e4ca-c776-3380-933b-f1218c4ef1fd (T-CD4+)
 hggraphid:a28e7725-1919-34f0-a648-45721d8bd6a2 (common cold)

This states that "SARS-CoV2" reactive "CD4+ T-cells" exist in ~40%-60% of unexposed individuals, suggesting cross-reactive T-cell recognition with "common cold"

responses to this virus.
 The first steps for such an understanding are the ability to quantify the virus-specific CD4⁺ and CD8⁺ T cells. Such knowledge is of immediate relevance, as it will provide insights into immunity and pathogenesis of SARS-CoV-2 infection, and the same knowledge will assist vaccine design and evaluation of candidate vaccines. Estimations of immunity are also central to epidemiological model calibration of future social distancing pandemic control measures (Kissler et al., 2020). Such projections are dramatically

baum,¹
 Jolla, CA
 92037, USA

This is a part of text that has a specific style (*italic*)

This is a result reported on in this paper

doi:10.1126/sciimmunol.aan5393

This is a citation of another paper

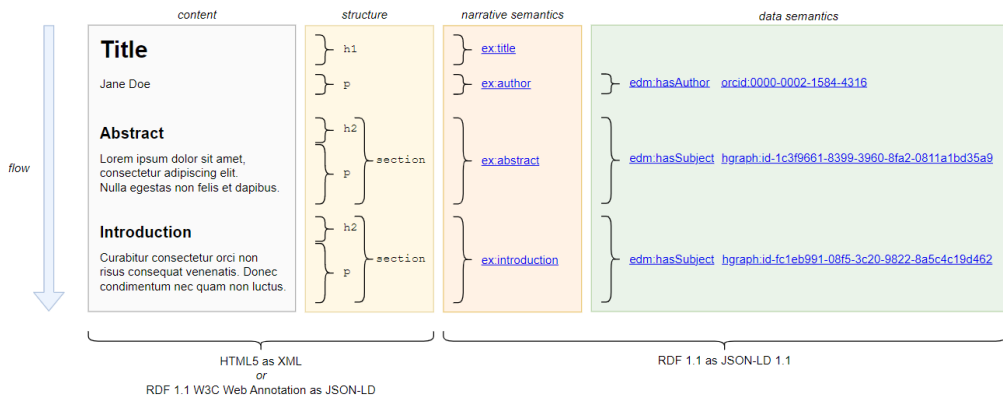
Cell 181, 1489–1501, June 25, 2020 © 2020 Elsevier Inc. 1489

Provenance standards: all together now!

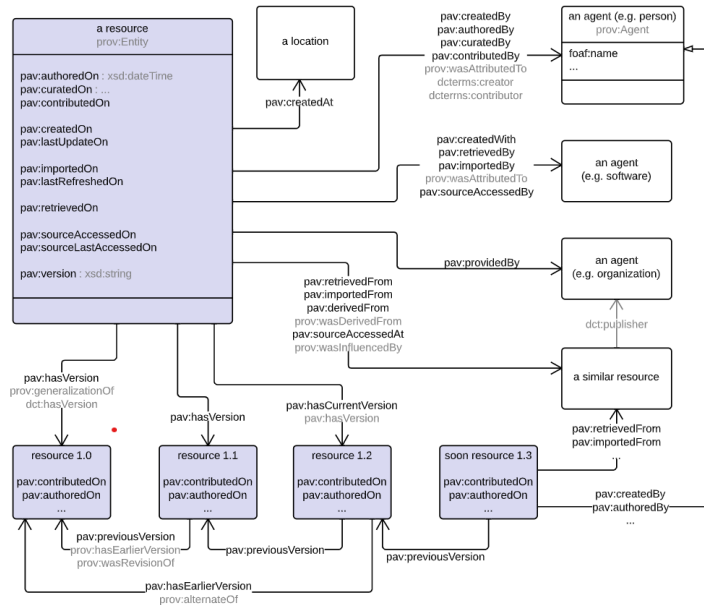
NISO Content Profile/Linked Document Group



W3C Provenance, Authoring and Versioning

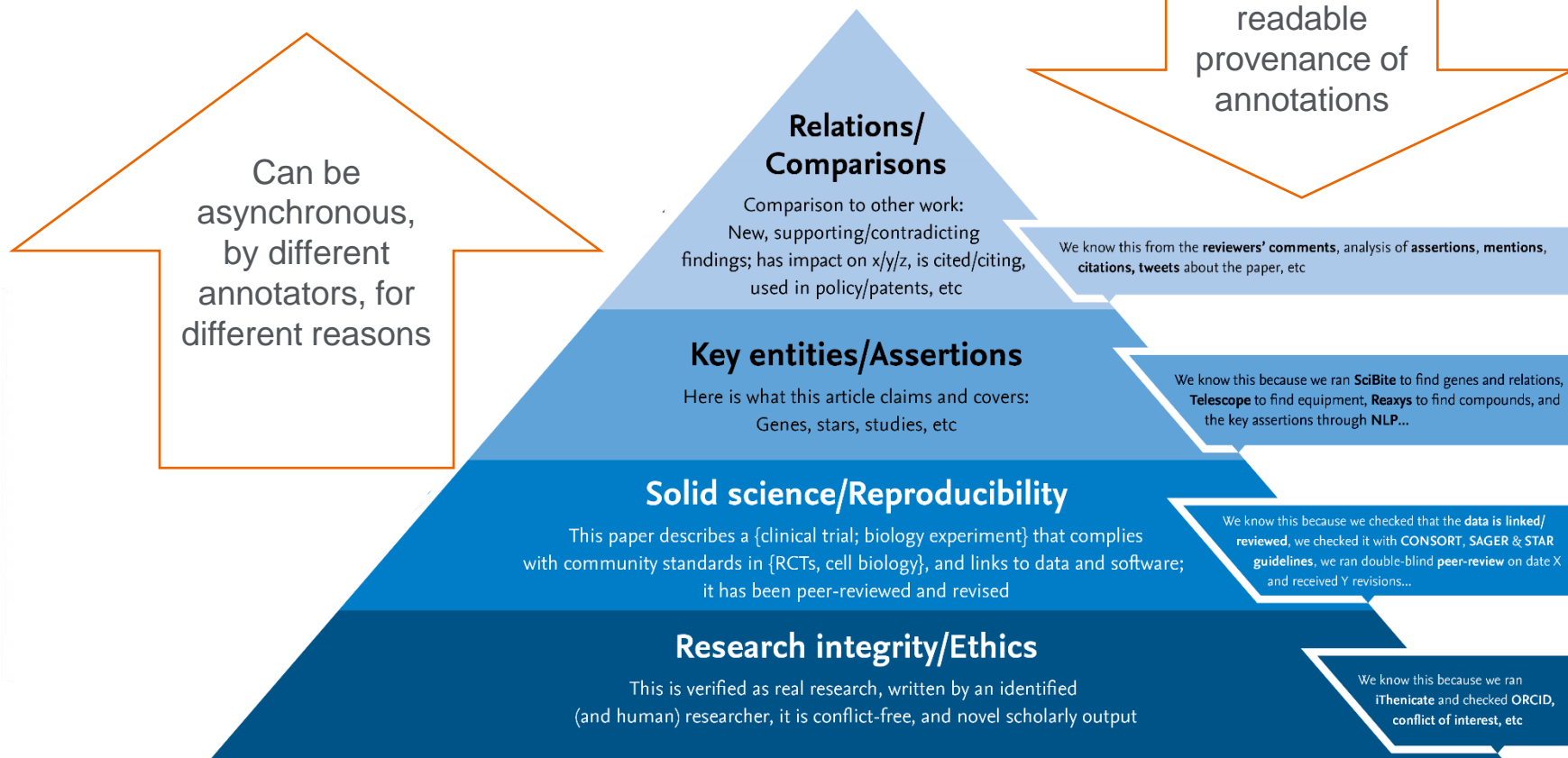


<https://www.niso.org/standards-committees/cpld>



<https://pav-ontology.github.io/pav/>

Building trust through provenance:



Anything else re.: “How will AI influence the work of libraries and librarians, especially in the humanities?”



Here are a few ways in which AI may impact libraries and librarians in the humanities:

1. **Discovery and Search:** AI-powered search and discovery tools can help users find relevant resources quickly and efficiently, even among large collections. AI can also be used to improve the quality of search results by using natural language processing (NLP) and machine learning algorithms to understand the context and intent of user queries.
2. **Personalization:** AI-powered personalization systems can be used to provide users with tailored recommendations of resources based on their research interests and browsing history. This can help increase the relevance and effectiveness of library resources for users.
3. **Content analysis and understanding:** AI can be used to analyze large amounts of unstructured data, including text, images, and audio. For example, NLP techniques can be used to automatically extract key concepts, entities, and relationships from texts, making it easier for librarians and researchers to analyze large collections of humanities materials such as books, journals, and manuscripts.
4. **Digital Humanities:** AI can also be used to support digital humanities research, such as text-mining, sentiment analysis and sentiment extraction, which can aid in research on large datasets of historical, literary and linguistic content.
5. **Automation:** AI can help automate repetitive tasks such as data entry, classification, and metadata creation. This can free up librarians' time to focus on more complex and value-adding tasks such as research support and instruction.
6. **Metadata generation:** With the help of machine learning, AI can help automate the process of metadata creation, speeding up the process and also providing more accurate descriptions of digital resources.
7. **Preservation and Conservation:** AI can help with the identification, preservation and conservation of digital heritage. For example, computer vision techniques can be used to help identify and preserve images and other media in digital collections.

Maybe ChatGPT has evolved in its thinking about this?



GPT3.5, 1/16/23

1. **Discovery and Search:** AI-powered search and discovery tools can help users find relevant resources quickly and efficiently, even among large collections,,,
2. **Personalization:** AI-powered personalization systems can be used to provide users with tailored recommendations of resources based on their research interests and browsing history,,,
3. **Content analysis and understanding:** AI can be used to analyze large amounts of unstructured data, including text, images, and audio,,,
4. **Digital Humanities:** AI can also be used to support digital humanities research, such as text-mining, sentiment analysis and sentiment extraction,,,
5. **Automation:** AI can help automate repetitive tasks such as data entry, classification, and metadata creation,,,
6. **Metadata generation:** With the help of machine learning, AI can help automate the process of metadata creation,,,
7. **Preservation and Conservation:** AI can help with the identification, preservation and conservation of digital heritage,,

GPT4, 4/28/23

1. **Cataloging and metadata generation:** AI can help automate cataloging processes and generate metadata for library resources. By using machine learning algorithms,,
2. **Semantic analysis and text mining:** AI-powered semantic analysis and text mining can help librarians better understand and organize vast collections of texts in the humanities...
3. **Improved search and discovery:** AI can enhance search functionality by understanding natural language queries, providing more accurate and relevant search results...
4. **Personalized recommendations:** AI can analyze user behavior, preferences, and search history to generate personalized reading recommendation...
5. **Enhanced digital collections:** AI can be used to transcribe, translate, or annotate digital resources, making them more accessible and engaging for users. ...
6. **Virtual research assistants:** AI-powered research assistants can help patrons with tasks like literature reviews, citation management, and data analysis....
7. **Digital humanities projects:** AI can be used in various digital humanities projects, such as text analysis, topic modeling, and sentiment analysis, to gain new insights into historical texts, cultural artifacts, and other humanities resources.
8. **Predictive analytics:** AI can help librarians identify trends and emerging research areas in the humanities, enabling them to make informed decisions about collection development and resource allocation.
9. **Skill development and training:** AI-driven tools can be used to create personalized learning experiences for librarians, helping them...

Um yeah. A relentless pace, indeed!

Similar (though not the same).

In January but not in April.

New in April.



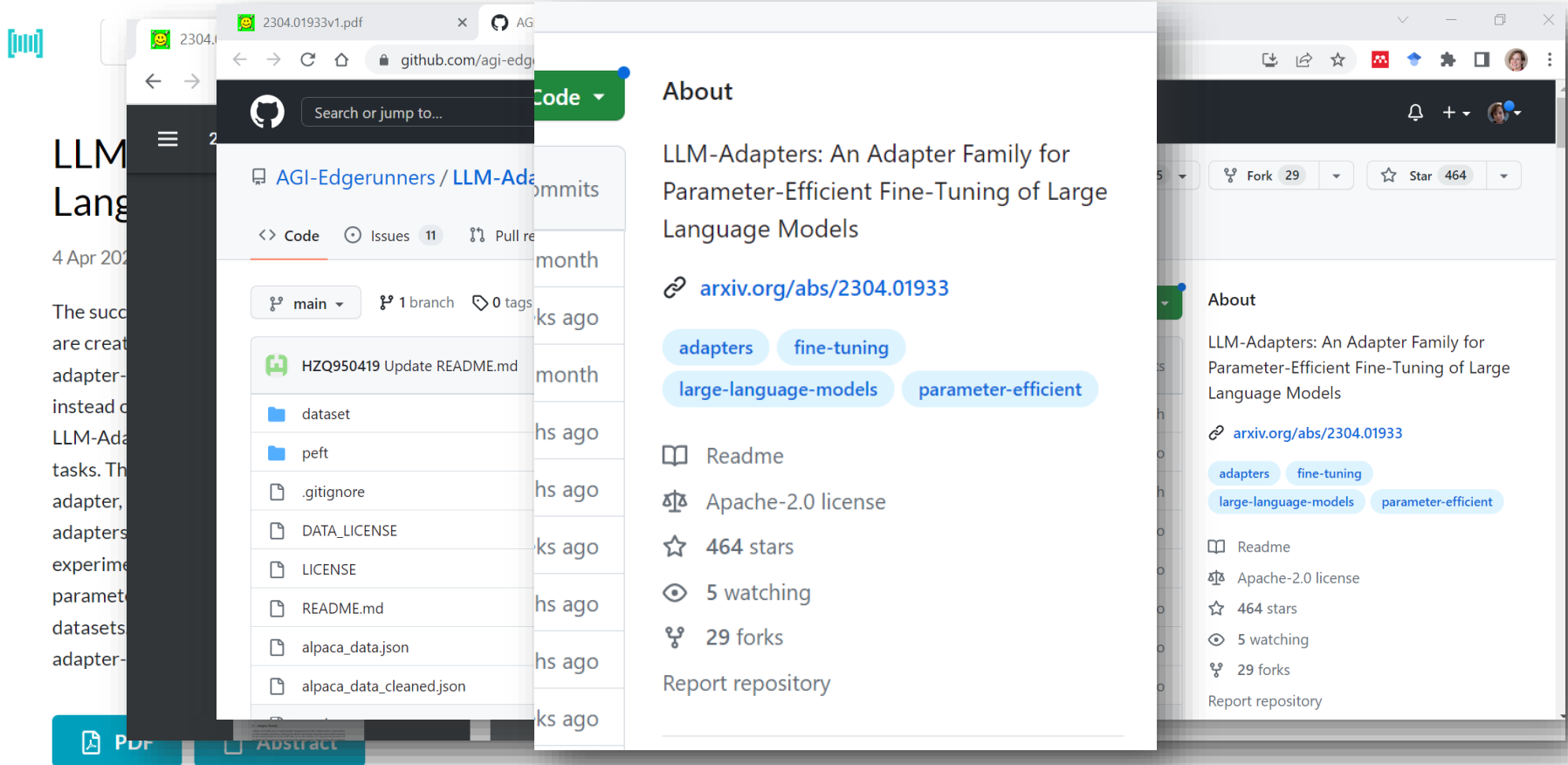
GPT3.5, 1/6/23

1. Discovery and Search: AI-powered search and discovery tools can help users find relevant resources quickly and efficiently, even among large collections,,,
2. Personalization: AI-powered personalization systems can be used to provide users with tailored recommendations of resources based on their research interests and browsing history,,,
3. Content analysis and understanding: AI can be used to analyze large amounts of unstructured data, including text, images, and audio,,,
4. Digital Humanities: AI can also be used to support digital humanities research, such as text-mining, sentiment analysis and sentiment extraction,,,
5. Automation: AI can help automate repetitive tasks such as data entry, classification, and metadata creation,,,
6. Metadata generation: With the help of machine learning, AI can help automate the process of metadata creation,,,
7. Preservation and Conservation: AI can help with the identification, preservation and conservation of digital heritage,,

GPT4, 4/28/23

1. Cataloging and metadata generation: AI can help automate cataloging processes and generate metadata for library resources. By using machine learning algorithms,,
2. Semantic analysis and text mining: AI-powered semantic analysis and text mining can help librarians better understand and organize vast collections of texts in the humanities...
3. Improved search and discovery: AI can enhance search functionality by understanding natural language queries, providing more accurate and relevant search results...
4. Personalized recommendations: AI can analyze user behavior, preferences, and search history to generate personalized reading recommendation...
5. Enhanced digital collections: AI can be used to transcribe, translate, or annotate digital resources, making them more accessible and engaging for users. ...
6. Virtual research assistants: AI-powered research assistants can help patrons with tasks like literature reviews, citation management, and data analysis....
7. Digital humanities projects: AI can be used in various digital humanities projects, such as text analysis, topic modeling, and sentiment analysis, to gain new insights into historical texts, cultural artifacts, and other humanities resources.
8. Predictive analytics: AI can help librarians identify trends and emerging research areas in the humanities, enabling them to make informed decisions about collection development and resource allocation.
9. Skill development and training: AI-driven tools can be used to create personalized learning experiences for librarians, helping them...

How is this research published? shared? used!



The image displays a collage of overlapping screenshots related to the LLM-Adapters research project. The central focus is the GitHub repository page for `AGI-Edgerunners / LLM-Adapters`. The repository page shows the title "LLM-Adapters: An Adapter Family for Parameter-Efficient Fine-Tuning of Large Language Models", the arXiv preprint link arxiv.org/abs/2304.01933, and several topic tags: `adapters`, `fine-tuning`, `large-language-models`, and `parameter-efficient`. It also displays repository statistics: 464 stars, 5 watching, and 29 forks. The repository content includes a README, Apache-2.0 license, and files like `dataset`, `peft`, `.gitignore`, `DATA_LICENSE`, `LICENSE`, `README.md`, `alpaca_data.json`, and `alpaca_data_cleaned.json`. Overlapping this is a screenshot of the arXiv preprint page, which mirrors the repository information. On the left, a PDF viewer shows a document titled "LLM Lang" with a date of "4 Apr 2023" and a snippet of text: "The success are created adapter- instead of LLM-Ada tasks. The adapter, adapters experime paramet datasets adapter-". At the bottom, there are buttons for "PDF" and "Abstract".

In Summary (1/16/23):

From fake news to papermills to plagiarizing chatbots, the knowledge objects that scholars, libraries and publishers, create, share and curate are subject to doubt. So how can we restore confidence in research? AI is here to stay, and will increasingly gain ground as a partner in the research process. But we need to make sure that what we create and curate together in this cyborgian knowledge environment serves humanity, and the future we hope to create. This talk will raise a few questions and share a few examples on how [we can use AI to combat the challenges of doubt and misinformation in research.](#)

[I'll discuss how AI can be used to verify sources, detect plagiarism and generate trustworthy research outputs. Additionally, we will explore the potential of AI to assist with data analysis, identify patterns and generate new insights. However, as we embrace AI in research, we must also be mindful of the ethical implications and potential biases that could arise. Ultimately, we need to work collaboratively to ensure that AI is used in a way that aligns with our values and contributes to a sustainable and equitable future for all. </>](#)

In Summary (5/1/23):

From fake news to papermills to plagiarizing chatbots, the knowledge objects that scholars, libraries and publishers, create, share and curate are subject to doubt. So how can we restore confidence in research? AI is here to stay, and will increasingly gain ground as a partner in the research process. But we need to make sure that what we create and curate together in this cyborgian knowledge environment serves humanity, and the future we hope to create. This talk will raise a few questions and share a few examples on how **<.>we can leverage the power of AI to enhance the credibility and integrity of research.**

By developing robust systems that effectively detect and mitigate the risks of misinformation, plagiarism, and other unethical practices, we can work towards fostering a more reliable and trustworthy knowledge ecosystem. [...]

GPT4 elaborates on what we need to do:



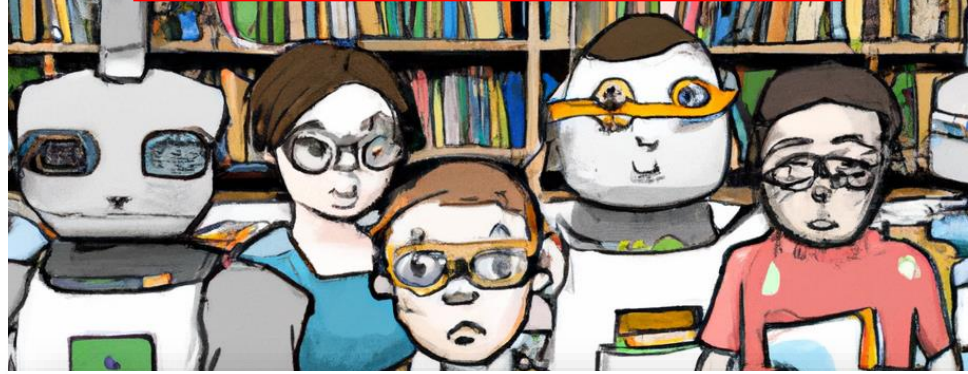
1. First, we must create **AI-driven tools that can effectively identify and filter out fake news** and misinformation. This can be achieved by training machine learning models on vast datasets of verified and fact-checked content, enabling them to discern factual information from falsehoods. Additionally, these tools should be designed to flag any suspicious patterns or inconsistencies, prompting further investigation from human experts.
2. Second, we **should implement stringent checks and balances to prevent paper mills** and plagiarism from undermining the quality of published research. This involves developing advanced algorithms that can detect similarities between submitted manuscripts and previously published work, as well as cross-referencing authorship data to uncover potential conflicts of interest.
3. Third, we must **focus on promoting transparency and open communication within the research community**. This can be accomplished by creating collaborative platforms that facilitate peer review, reproducibility, and data sharing among researchers. AI can play a critical role in streamlining these processes, making it easier for scholars to collaborate and engage in rigorous scholarly discourse.
4. Lastly, we should **invest in educating researchers, librarians, and publishers on the ethical use of AI** in the research process. This includes providing resources and training on best practices for data management, algorithmic bias, and the responsible use of AI-generated content. [...]

Question that remains:

By addressing these challenges, **we** can harness the full potential of AI to augment human intelligence, ensuring that the knowledge **we** create and share is reliable, trustworthy, and serves the greater good of humanity.

As **we** move forward in this era of AI-assisted research, **we** must remain vigilant and dedicated to preserving the integrity of **our shared** knowledge, fostering a future where truth and innovation flourish.

Who are “we”??



The Writers Guild of America said it wants Hollywood’s top studios and networks to regulate the use of AI on creative projects. The union’s specific demand, according to a document released Monday, states: “AI can’t write or rewrite literary material; can’t be used as source material; and MBA-covered [contract-covered] material can’t be used to train AI.”

