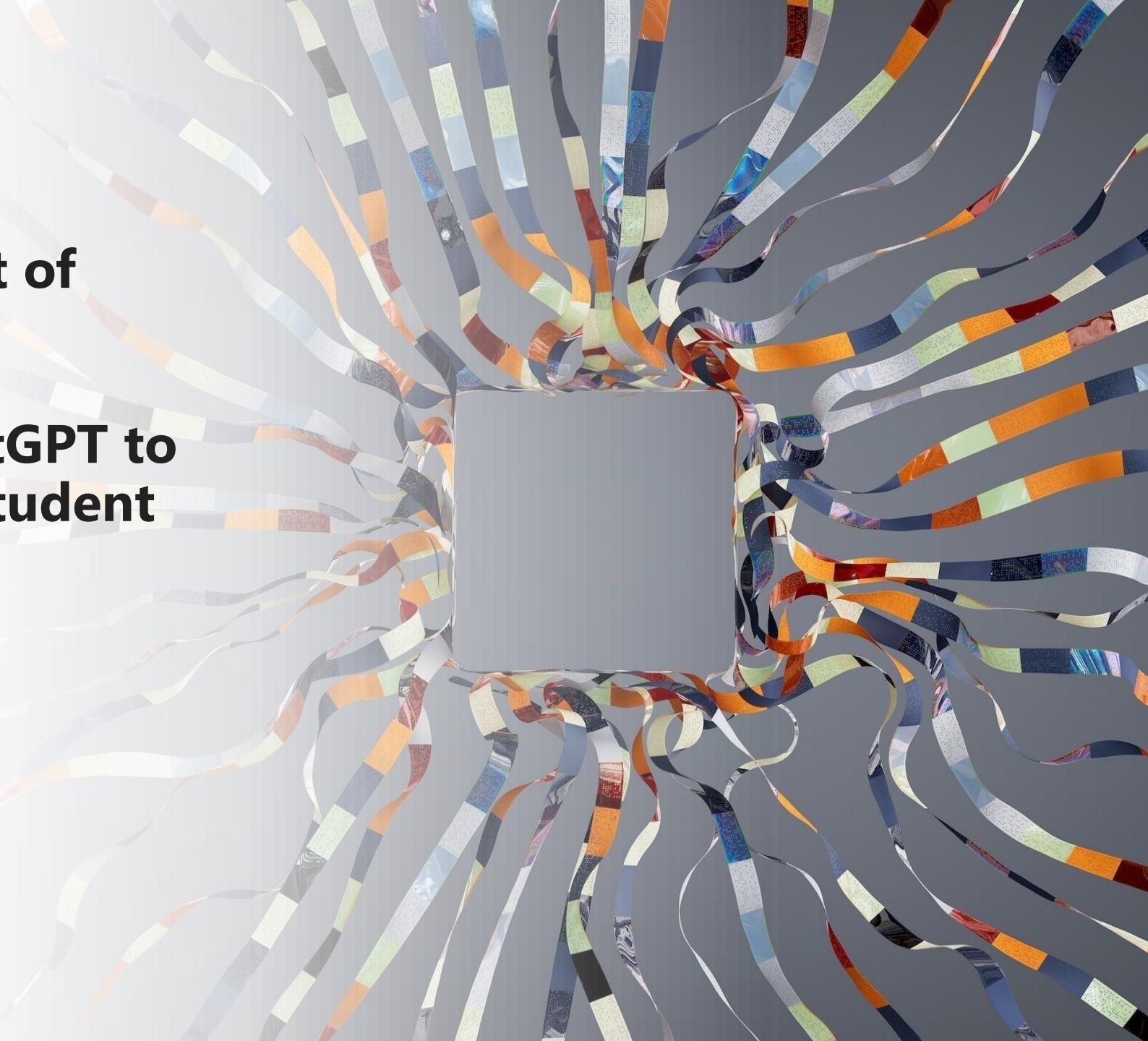




AI literacy in the context of working with sources: Identifying and testing strategies for using ChatGPT to find references from a student perspective

Dr. Tine Wirenfeldt Jensen, University of Southern Denmark/METoDo, Denmark

Søren Wirenfeldt Jensen, IttyBits, Denmark



- PhD in Learning & Education
- Worked with academic writing and academic literacy for 20 years
- Member of the Danish Ministry of Education's expert group on ChatGPT (Generative AI)
- Member of an advisory board on ChatGPT/Generative AI for a Danish University
- Jensen, T. W. 2023. "Wiping away the stain: From "plagiarism checkers" to AI-assisted "plagiarism removers" and the implications for students' understanding of academic writing". *European Association for the Teaching of Academic Writing conference, Writing technology, thinking, and learning, Schweiz 2023.*
- Jensen, T. W. & S.W. Jensen. 2023. "AI literacy in the context of working with sources: Pitfalls and possibilities of AI-based natural language production systems in academic writing". *European Association for the Teaching of Academic Writing conference, Writing technology, thinking, and learning, Schweiz 2023*



The rise of OpenAI's ChatGPT

"Coming out of nowhere at the end of November 2022, ChatGPT had about 266 million visits by December – or about the same as Yahoo News, more than for the vast majority of websites – and ended April 2023 with about 1.76 billion visits, according to preliminary Similarweb estimates".

Source: <https://www.similarweb.com/blog/insights/ai-news/>

Students' use of ChatGPT

"5,894 students from across Swedish universities were surveyed about their use of and attitudes towards AI for learning purposes, both about chatbots (such ChatGPT) and other AI language tools (such as Grammarly)"

"Almost all the respondents are familiar with ChatGPT (but typically not with other chatbots); more than a third use ChatGPT regularly."

Malmström, H., Stöhr, C., & Ou, A. W. (2023). Chatbots and other AI for learning: A survey of use and views among university students in Sweden. (*Chalmers Studies in Communication and Learning in Higher Education* 2023:1)
<https://doi.org/10.17196/cls.csclhe/2023/0>

An increasing range of digital interfaces

“Students are faced with an increasing range of digital platforms with which to work, and an often unpredictable set of social and material resources which shape their writing and knowledge production.”

Bhatt, I., & MacKenzie, A. (2019). Just Google it! Digital literacy and the epistemology of ignorance. *Teaching in Higher Education*, 24(3), 302-317.

- Google
- Google Scholar
- Library databases
- SciSpace
- Mendeley
- Elicit
- ResearchRabbit
- Scite.ai
- ChatGPT

The need to be critically aware and reflective

“If students are restricted in what they can know because they are unaware of how exogenous actors (e.g. algorithms) actually work, and how they guide their choices and shape their experiences online, then it becomes important to educate them to be critically aware during their digital searches for information, research and critical argument, and to educate them to be reflective about their ritualized practices with digital literacy.”

Bhatt, I., & MacKenzie, A. (2019). Just Google it! Digital literacy and the epistemology of ignorance. *Teaching in Higher Education*, 24(3), 302-317.

A new type of search output: The hallucinated reference

ChatGPT can generate “plausible-looking but fake references” (Sharples, 2022).

“Language models (LMs) famously hallucinate, meaning that they fabricate strings of plausible but unfounded text. As LMs become more accurate, their fabrications become more believable and therefore more problematic. A primary example is “hallucinated references” to non-existent articles with titles readily fabricated by the LM.”

(Agrawal, et al., 2023).

(See Day, 2023; Gravel et al., 2023; for discussion within a disciplinary field)

Can ChatGPT produce "useable" references?

- Goal: Generate correct title and author
- Testing across a topic in three fields
 - Rhetoric: Racial Discourse and Language
 - Biology: Cell-free systems
 - Law: Privacy Rights and Surveillance
- Six strategies/prompts
 - 1 – 3: with the sole aim of generating references
 - 4 – 6: references were generated in the context of a literature review

Method

- Every prompt was executed three times, per topic, per strategy
- ChatGPT-3.5 generated 344 references in total (Spring 2023)
- ChatGPT-4 generated 646 references with the same prompts (July 2023)
- ChatGTP was asked to convert references to BibTeX-format
- For each reference, a programmatic search was conducted on Google Scholar to retrieve the top result
- The titles and authors of each reference from ChatGPT and Google Scholar were programmatically compared using text comparison tool using a fuzzy algorithm
- Misses were checked manually

Red is incorrect

References:

Church, G. M. (2010). Genome-scale engineering for systems and synthetic biology.

Proceedings of the National Academy of Sciences, 107(50), 21,167-21,172.

Endy, D. (2005). Foundations for engineering biology. *Nature*, 438(7067), 449-453.

Jewett, M. C., & Swartz, J. R. (2008). Cell-free biosynthesis of therapeutic proteins. *Current Opinion in Biotechnology*, 19(5), 450-457.

Noireaux, V., & Libchaber, A. (2008). Genetic circuit building blocks for cellular computation, communications, and signal processing. *Natural Computing*, 7(3), 317-339.

Quake, S. R., & Scherer, A. (2008). From micro- to nanofabrication with soft materials. *Science*, 322(5907), 1500-1504.

Swartz, J. R. (2006). Cell-free protein synthesis: An emerging technology paradigm for bottom-up proteomics. *Trends in Biotechnology*, 24(4), 161-166.

TI

List six references on the topic "Racial Discourse and Language" with a high number of Google Scholar citations

< 1 / 3 >



Sure! Here are six highly cited references on the topic of "Racial Discourse and Language" in no particular order:

1. Bonilla-Silva, E. (2003). *Racism without racists: Color-blind racism and the persistence of racial inequality in America*. Rowman & Littlefield Publishers.
2. van Dijk, T. A. (1993). *Elite discourse and racism*. Sage.
3. Essed, P. (1991). *Understanding everyday racism: An interdisciplinary theory*. Sage.
4. Hill, J. H. (2008). Language, race, and white public space. *American Anthropologist*, 110(4), 507-518.
5. Smitherman, G. (1994). *Talkin and testifyin: The language of black America*. Wayne State University Press.
6. Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist*, 62(4), 271-286.

I hope this helps!

Six strategies (student perspective)

1. Make reference
2. Don't lie – verifiable links
3. Highly cited
4. In the tone of a professor
5. Conversational style, using the chat interface
6. Ask ChatGPT to generate the prompt – adjusted version for references

Six strategies (student perspective)

1. Make reference
2. Don't lie – verifiable links
3. Highly cited
4. In the tone of a professor
5. Conversational style, using the chat interface
6. Ask ChatGPT to generate the prompt – adjusted version for references

1. Make references

Prompt:

"List six references on the topic {insert topic}"

1. Make references: Results

	Pct of useable references	
	GPT-3.5	GPT-4
Biology: Cell-free systems	56%	72%
Law: Privacy rights	83%	100%
Rethoric: Racial Discourse and Language	78%	89%

2. Don't lie

Prompt:

"List six references on the topic "{insert topic}". Ensure that all sources included in the review are non-fictitious and have links for verification".

2. Don't lie: Results

	Pct of useable references	
	GPT-3.5	GPT-4
Biology: Cell-free systems	67%	28%
Law: Privacy rights	6%	89%
Rethoric: Racial Discourse and Language	50%	100%

Titles with no author – not a match in our setup

The screenshot illustrates a workflow for managing bibliographic entries. On the left, a dark-themed BibTeX editor displays two entries:

```
bibtex
@article{biotech_advances,
    title={Cell-Free Protein Synthesis: Applications Come of Age},
    author={Unknown},
    journal={Biotechnology Advances},
    volume={Unknown},
    pages={Unknown},
    year={Unknown}
}

@article{frontiers_chem,
    title={Cell-free systems in the new age of synthetic biology},
    author={Unknown},
    journal={Frontiers in Chemistry},
    volume={Unknown},
    pages={Unknown},
    year={Unknown}
}
```

The 'author' field for both entries is highlighted with a red box. On the right, a search result from Google Scholar is shown for the first entry. The title 'Cell-free protein synthesis: applications come of age' is highlighted with a green box. The result includes the authors (ED Carlson, R Gan, CE Hodgman, MC Jewett), the journal ('Biotechnology advances'), and a brief abstract. Below the abstract, citation statistics are provided: 99 citations, 741 cited by, and 14 versions.

3. Highly cited

Prompt:

"List six references on the topic "{insert topic}" with a high number of Google Scholar citations".

3. Highly cited: Results

	Pct of useable references	
	GPT-3.5	GPT-4
Biology: Cell-free systems	72%	100%
Law: Privacy rights	61%	94%
Rethoric: Racial Discourse and Language	100%	94%

4. In the tone of a professor

Prompt:

"Create a literature review in the tone of a professor for an academic paper on the topic of "{insert topic}". References should be in APA format"

4. In the tone of a professor: Results

	Pct of useable references	
	GPT-3.5	GPT-4
Biology: Cell-free systems	47%	60%
Law: Privacy rights	83%	89%
Rethoric: Racial Discourse and Language	86%	96%

5. Conversational style

Prompts:

- “List six influential scholars on the topic of “{insert topic}” in the field of {insert field}”
- “List a highly cited paper by each scholar”
- “Create a literature review using these papers as references. References should be in APA format”

5. Conversational style: Results

	Pct of useable references	
	GPT-3.5	GPT-4
Biology: Cell-free systems	22%	94%
Law: Privacy rights	89%	89%
Rethoric: Racial Discourse and Language	83%	100%

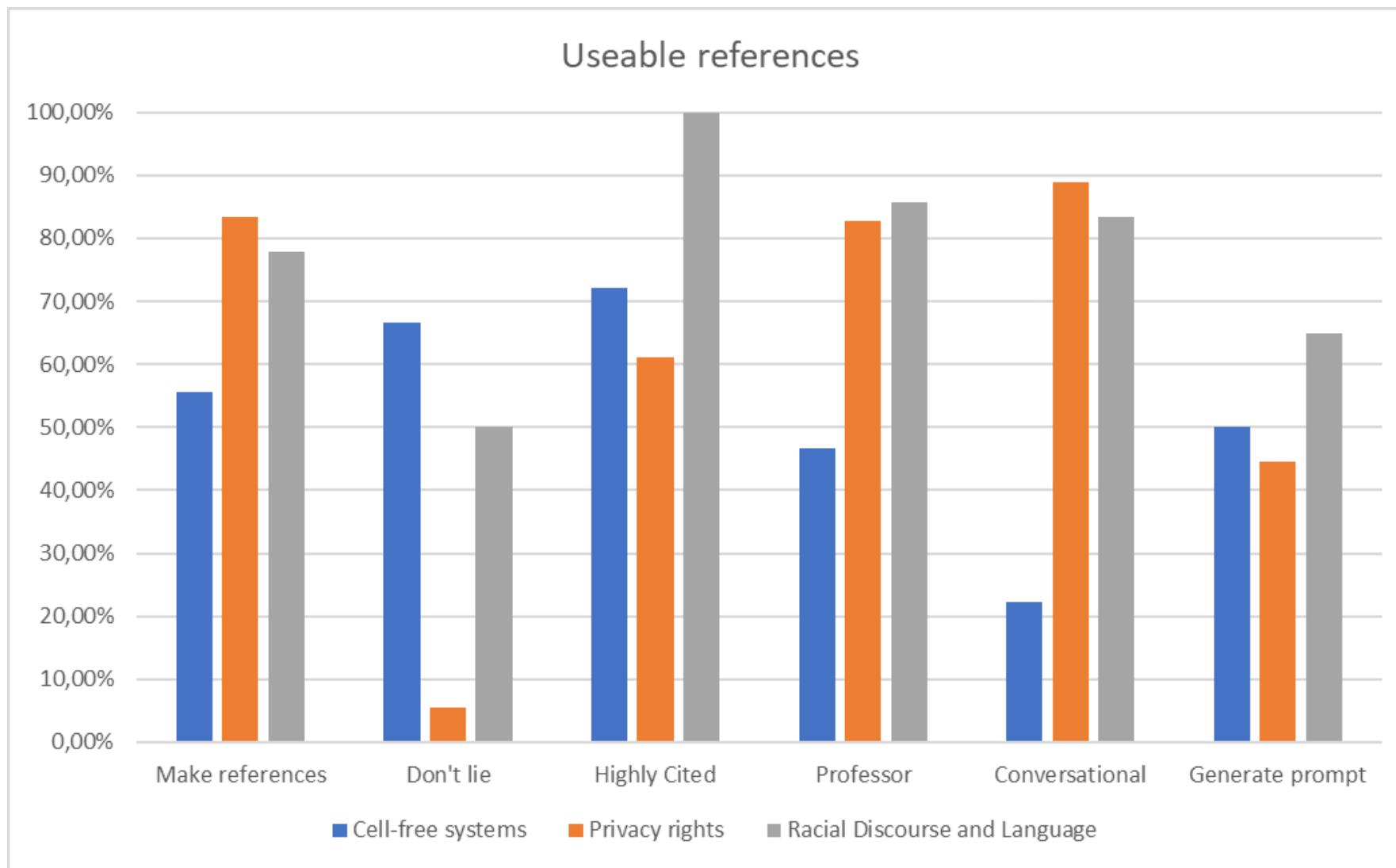
6. Ask ChatGPT to generate the prompt – adjusted version for references

“Write a literature review exploring the current research on {insert topic}. Your review should analyze and synthesize at least 6 relevant academic sources, highlighting the key themes, controversies, and gaps in knowledge in the field. Be sure to address the following questions: What are the main arguments and evidence presented by researchers in this field? What are the gaps and limitations in current research, and what future directions should researchers take? Finally, how does your review contribute to a deeper understanding of this field and its implications for {insert topic}? Use inline references in APA format, and be sure to add the references section”

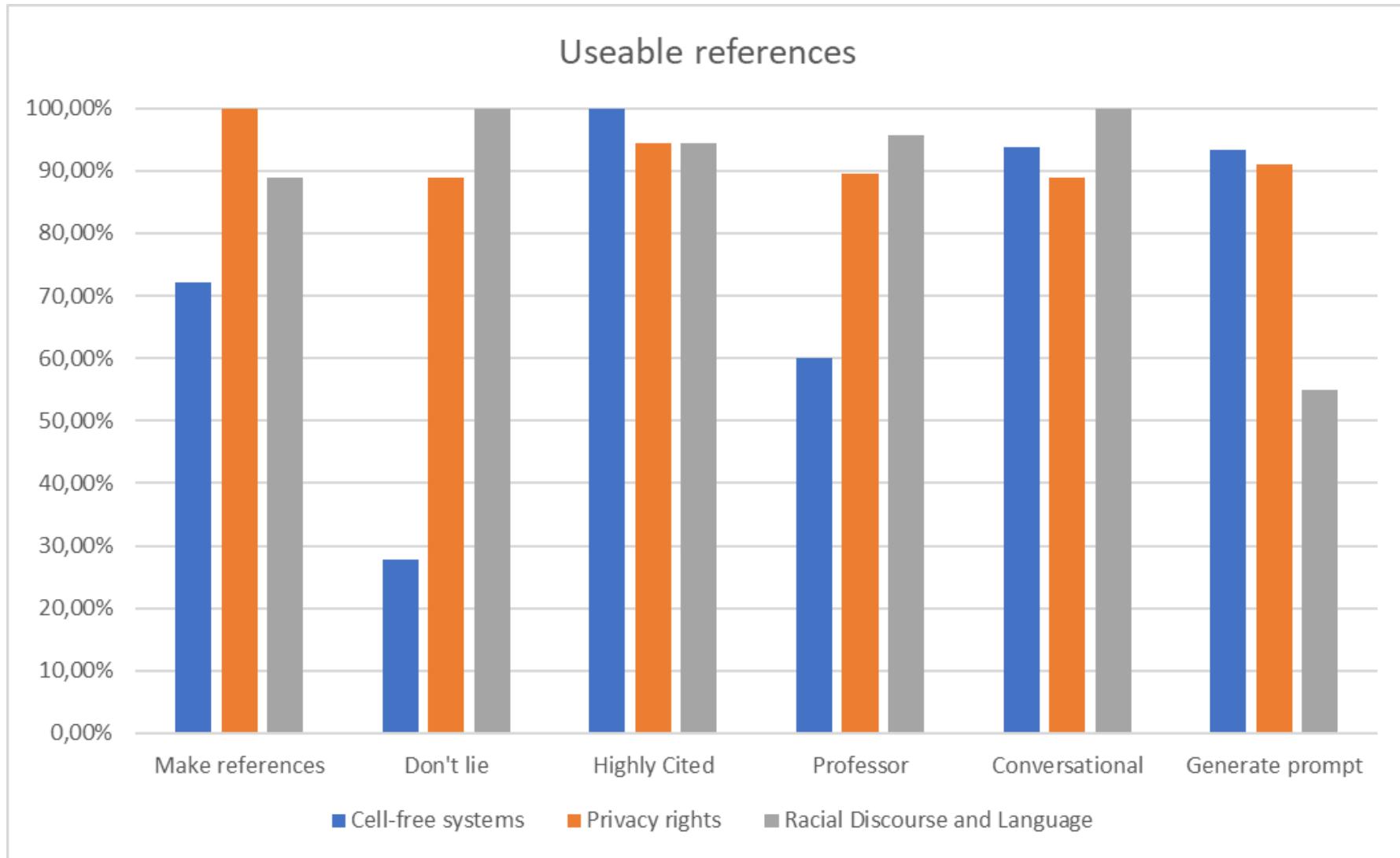
6. Ask ChatGPT to generate the prompt: Results

	Pct of useable references	
	GPT-3.5	GPT-4
Biology: Cell-free systems	50%	93%
Law: Privacy rights	44%	91%
Rethoric: Racial Discourse and Language	65%	55%

ChatGPT 3.5 (free version)



ChatGPT 4 (paid version)



Reinforcement learning from human preferences

I'm sorry for the confusion, but as an AI model developed by OpenAI, I don't have the ability to browse the internet or access real-time data or hyperlinks. However, I can provide examples of credible sources that you might look for, or ones that are likely to exist based on my training data up until September 2021. Please check their availability and content independently.

Remember, I can't guarantee the availability of these resources online, and I recommend checking a library or academic database.

For more on Reinforcement learning from human preferences see e.g. Menick et al. (2022). Teaching language models to support answers with verified quotes. *arXiv preprint arXiv:2203.11147*.

ChatGPT plugins beta (only in GPT-4)

The screenshot shows the plugin selection interface for ChatGPT. At the top, there are two tabs: "GPT-3.5" and "GPT-4". The "GPT-4" tab is highlighted with a purple puzzle-piece icon and is enclosed in a rounded rectangle. Below the tabs are three small icons: a stylized owl, a globe, and a red 'X'. A dropdown arrow is positioned next to the 'X' icon. The main area displays a list of four plugins, each with a small icon and a checkmark in a blue box:

- Scholar Assist (owl icon)
- ScholarAI (globe icon)
- Xpapers (red 'X' icon)

At the bottom of the list is a button labeled "Plugin store" with a right-pointing arrow.

Navigating a new complexity

- Both students and academic writing educators need to build their AI literacy (Long & Magerko, 2020; Laupichler et al., 2022) in the context of working with sources.
- In order to do this, we must not only be aware of how ChatGPT functions on a technical level, but also how it might perform from a student-user perspective.

References

- Agrawal, A., Mackey, L., & Kalai, A. T. (2023). Do Language Models Know When They're Hallucinating References?. *arXiv preprint arXiv:2305.18248*
- Anon. (2021). What grades can AI get in college?, Best Universities. Retrieved from <https://best-universities.net/features/what-grades-can-ai-get-in-college/>
- Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. Available at SSRN 4308687.
- Bhatt, I., & MacKenzie, A. (2019). Just Google it! Digital literacy and the epistemology of ignorance. *Teaching in Higher Education*, 24(3), 302-317.
- Day, T. (2023). A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT. *The Professional Geographer*, 1-4.

References

- Gravel, J., D'Amours-Gravel, M., & Osmanlliu, E. (2023). Learning to fake it: limited responses and fabricated references provided by ChatGPT for medical questions. *medRxiv*, 2023-03.
- Laupichler, M. C., Astar, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 100101
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI conference on human factors in computing systems (pp. 1-16).
- Menick et al. (2022). Teaching language models to support answers with verified quotes. *arXiv preprint arXiv:2203.11147*.
- Mishra, S., Khashabi, D., Baral, C., Choi, Y., & Hajishirzi, H. (2021). Reframing Instructional Prompts to GPTk's Language. *arXiv preprint arXiv:2109.07830*.
- Sharples, M. (2022). Automated essay writing: an AIED opinion. *International Journal of Artificial Intelligence in Education*, 32(4), 1119-1126.