BEYOND THE HYPE

Deploying and Evaluating a Conversational Agent Using LLMs in an Academic Setting

OBJECTIVES

Implement a retrieval-augmented generation (RAG) based system capable of answering reference questions

Develop an evaluation instrument and protocol to measure the “usefulness” of the chatbot and compare multiple models

TEAM

Megan Fitzgibbons (Librarian)

Joshua Chalifour (Librarian)

Olivier Charbonneau (Librarian)

Aviva Majerczyk (Research Assistant)

Yara Stouhi (Research Assistant)

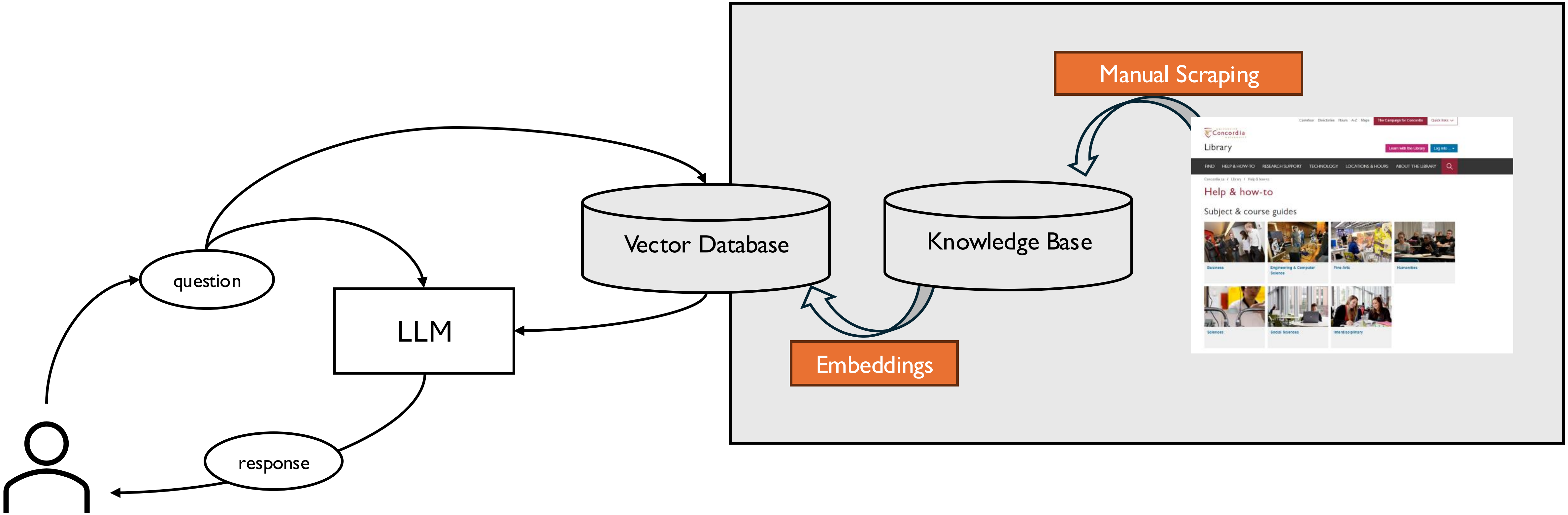
Francisco Berrizbeitia (Developer)

FUNDERS AND PARTNERS

Concordia Library Research Grant
(2023-2024)

Concordia Applied AI Institute:
Collaborations with Industry Grant
(2024-2025)

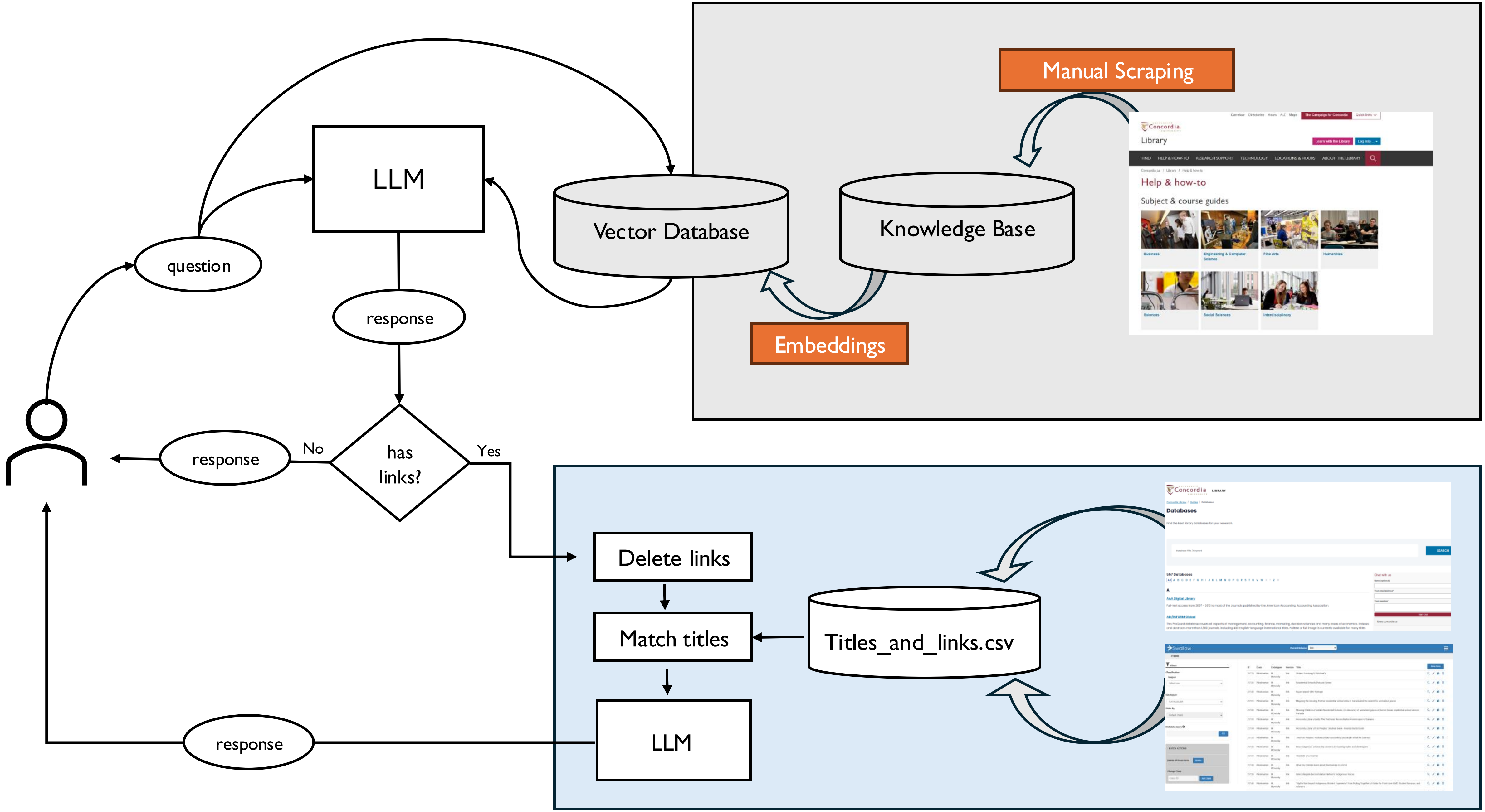
Cinémathèque Québécoise



Help & how-to

Subject & course guides

- Business
- Engineering & Computer Science
- Fine Arts
- Humanities
- Sciences
- Social Sciences
- Interdisciplinary


LLM
question
response
has links?
No
Yes
response
Vector Database
Knowledge Base
Manual Scraping
Embeddings
Carrefour Directories Hours A-Z Maps The Campaign for Concordia Quick links
Concordia
Library
Learn with the Library
Log into ...
FIND HELP & HOW-TO RESEARCH SUPPORT TECHNOLOGY LOCATIONS & HOURS ABOUT THE LIBRARY
Concordia.ca / Library / Help & how-to
Help & how-to
Subject & course guides
Business
Engineering & Computer Science
Fine Arts
Humanities
Sciences
Social Sciences
Interdisciplinary
Delete links
Match titles
LLM
Titles_and_links.csv
Databases
Swallow
response

EXPERIMENTAL DESIGN

Create a questionnaire consisting of commonly asked reference questions.

Define a rubric to grade the answers.

Run the evaluation questionnaire using three different LLMs with the same prompt.

- OpenAI ChatGPT turbo 3.5
- Gemini
- PHI-3

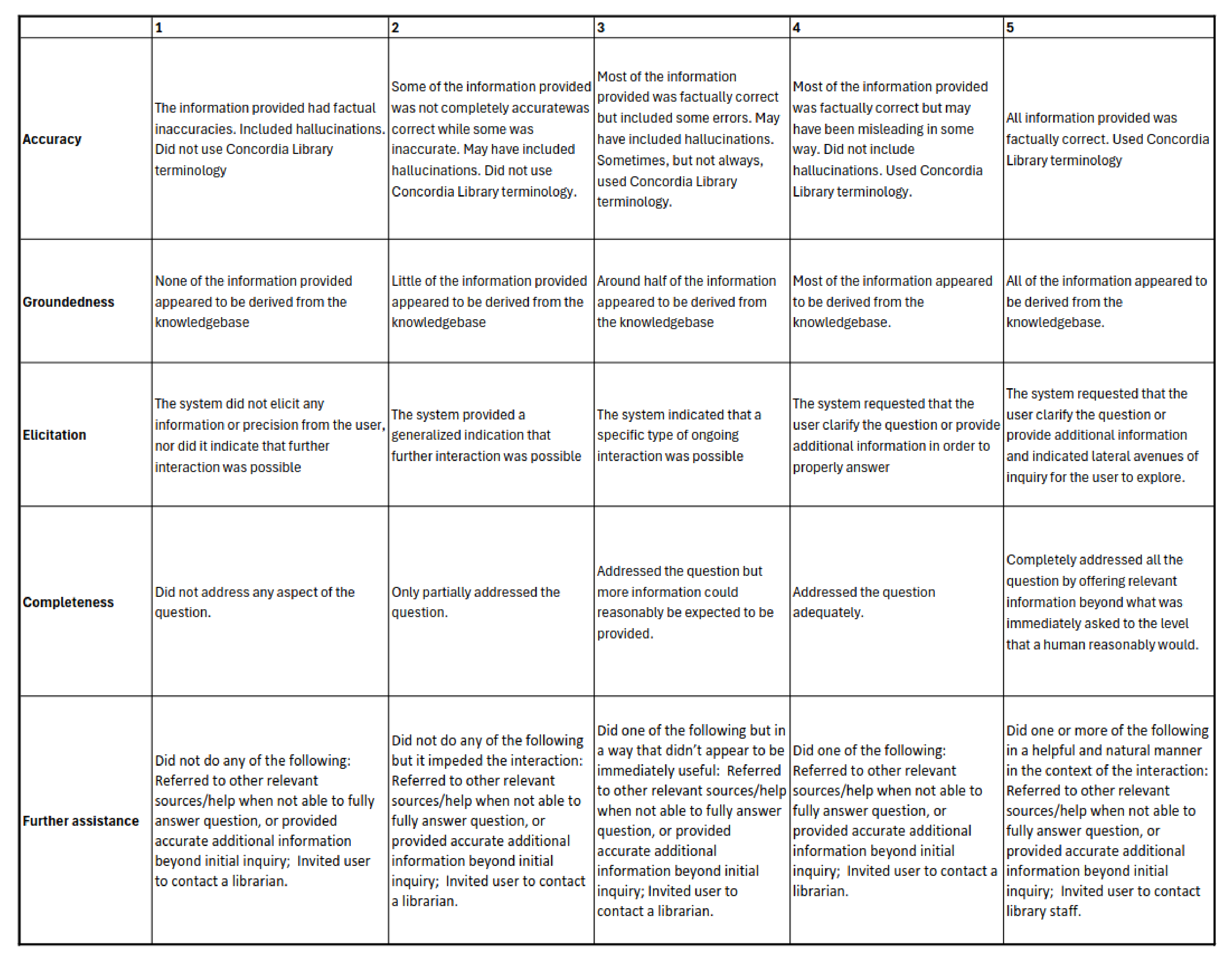
Grade all answers using the rubric (3 librarians and 1 student).

Discussion on the grading process.

INTERACTION PROMPTS

1. What should I do if I have a link and it's broken
2. Can I do an online class at the library?
3. How do I know if an article is peer-reviewed?
4. Can I rent textbooks?
5. How can I find primary sources?
6. Can I show a film in my class
7. Can I include an image from a website in my thesis
8. I have a research essay and don't know where to start
9. How do I request a book?
10. What if I need a book that Concordia doesn't have?
11. How can I download an eBook?
12. How can I find articles about social media methodology
13. How do I cite a source that I found referenced in another work?
14. Can you give me a link to a database for articles on the effects of climate change?

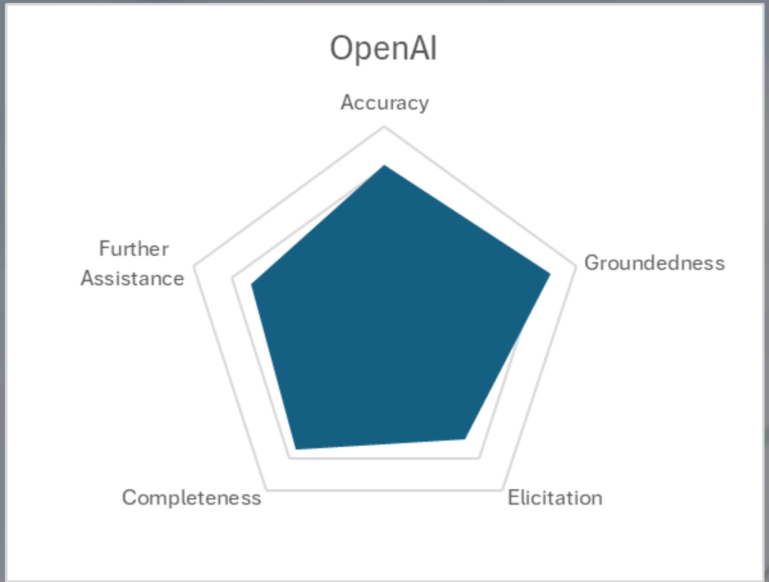
RUBRIC

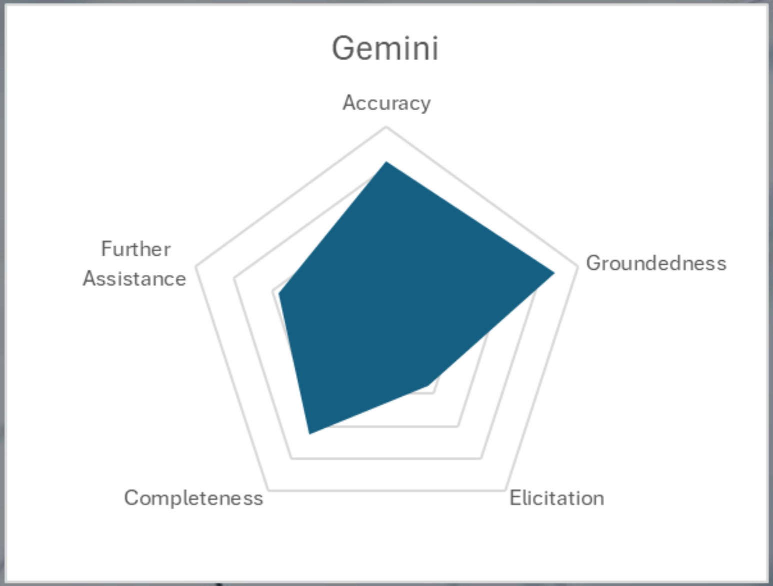


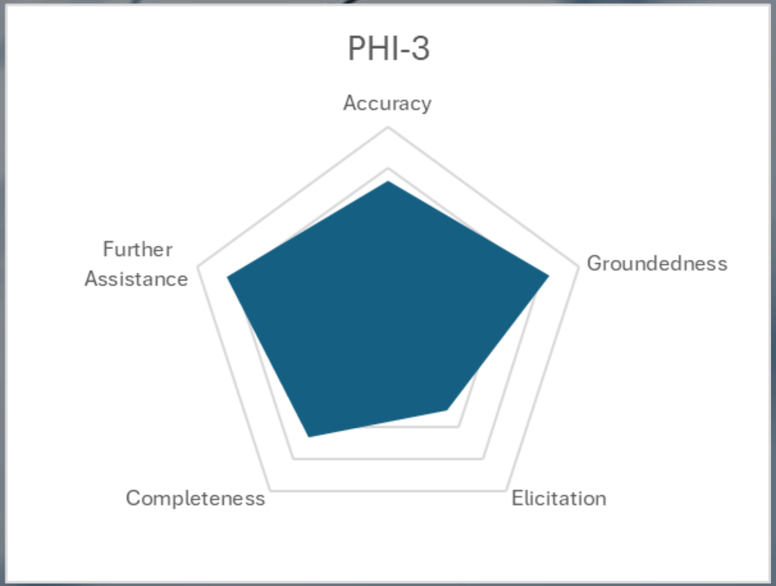
	1	2	3	4	5
Accuracy	The information provided had factual inaccuracies. Included hallucinations. Did not use Concordia Library terminology	Some of the information provided was not completely accuratewas correct while some was inaccurate. May have included hallucinations. Did not use Concordia Library terminology.	Most of the information provided was factually correct but included some errors. May have included hallucinations. Sometimes, but not always, used Concordia Library terminology.	Most of the information provided was factually correct but may have been misleading in some way. Did not include hallucinations. Used Concordia Library terminology.	All information provided was factually correct. Used Concordia Library terminology
Groundedness	None of the information provided appeared to be derived from the knowledgebase	Little of the information provided appeared to be derived from the knowledgebase	Around half of the information appeared to be derived from the knowledgebase	Most of the information appeared to be derived from the knowledgebase.	All of the information appeared to be derived from the knowledgebase.
Elicitation	The system did not elicit any information or precision from the user, nor did it indicate that further interaction was possible	The system provided a generalized indication that further interaction was possible	The system indicated that a specific type of ongoing interaction was possible	The system requested that the user clarify the question or provide additional information in order to properly answer	The system requested that the user clarify the question or provide additional information and indicated lateral avenues of inquiry for the user to explore.
Completeness	Did not address any aspect of the question.	Only partially addressed the question.	Addressed the question but more information could reasonably be expected to be provided.	Addressed the question adequately.	Completely addressed all the question by offering relevant information beyond what was immediately asked to the level that a human reasonably would.
Further assistance	Did not do any of the following: Referred to other relevant sources/help when not able to fully answer question, or provided accurate additional information beyond initial inquiry; Invited user to contact a librarian.	Did not do any of the following but it impeded the interaction: Referred to other relevant sources/help when not able to fully answer question, or provided accurate additional information beyond initial inquiry; Invited user to contact a librarian.	Did one of the following but in a way that didn't appear to be immediately useful: Referred to other relevant sources/help when not able to fully answer question, or provided accurate additional information beyond initial inquiry; Invited user to contact a librarian.	Did one of the following: Referred to other relevant sources/help when not able to fully answer question, or provided accurate additional information beyond initial inquiry; Invited user to contact a librarian.	Did one or more of the following in a helpful and natural manner in the context of the interaction: Referred to other relevant sources/help when not able to fully answer question, or provided accurate additional information beyond initial inquiry; Invited user to contact library staff.

Adapted from Lai, 2023

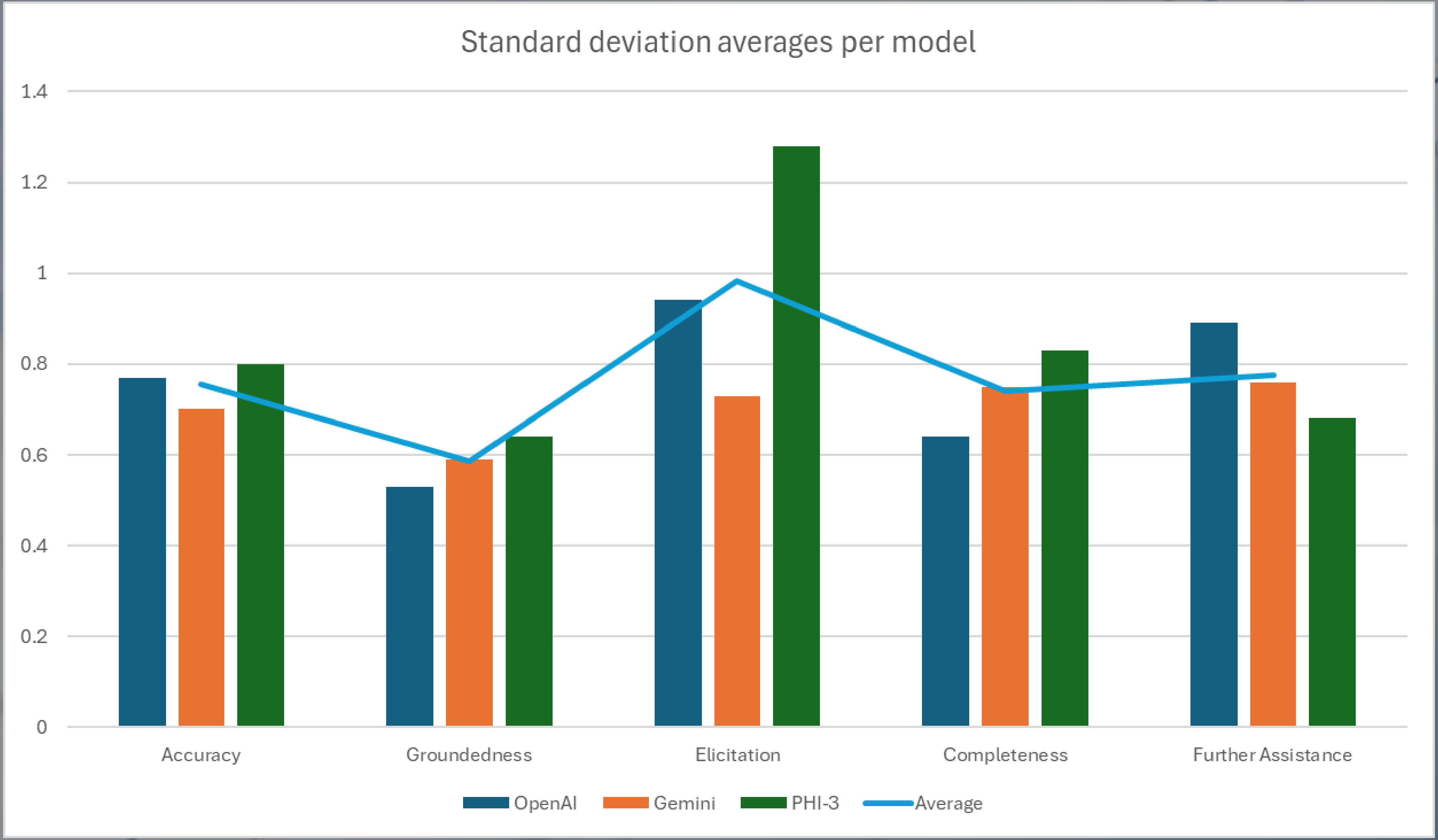
MODEL EVALUATION







RUBRIC EVALUATION



CONCLUSIONS

- RAG implementation requires upkeep
- Smaller language models might work as well as large ones
- Evaluation of performance is inherently subjective
- Protocol for testing requires iterations
- Raised questions about how we determine what value new tech provides to reference processes

POTENTIAL NEXT STEPS

- Change variable
- Fine tune Phi3
- Test a wider array of questions with revised rubric
- End-user testing

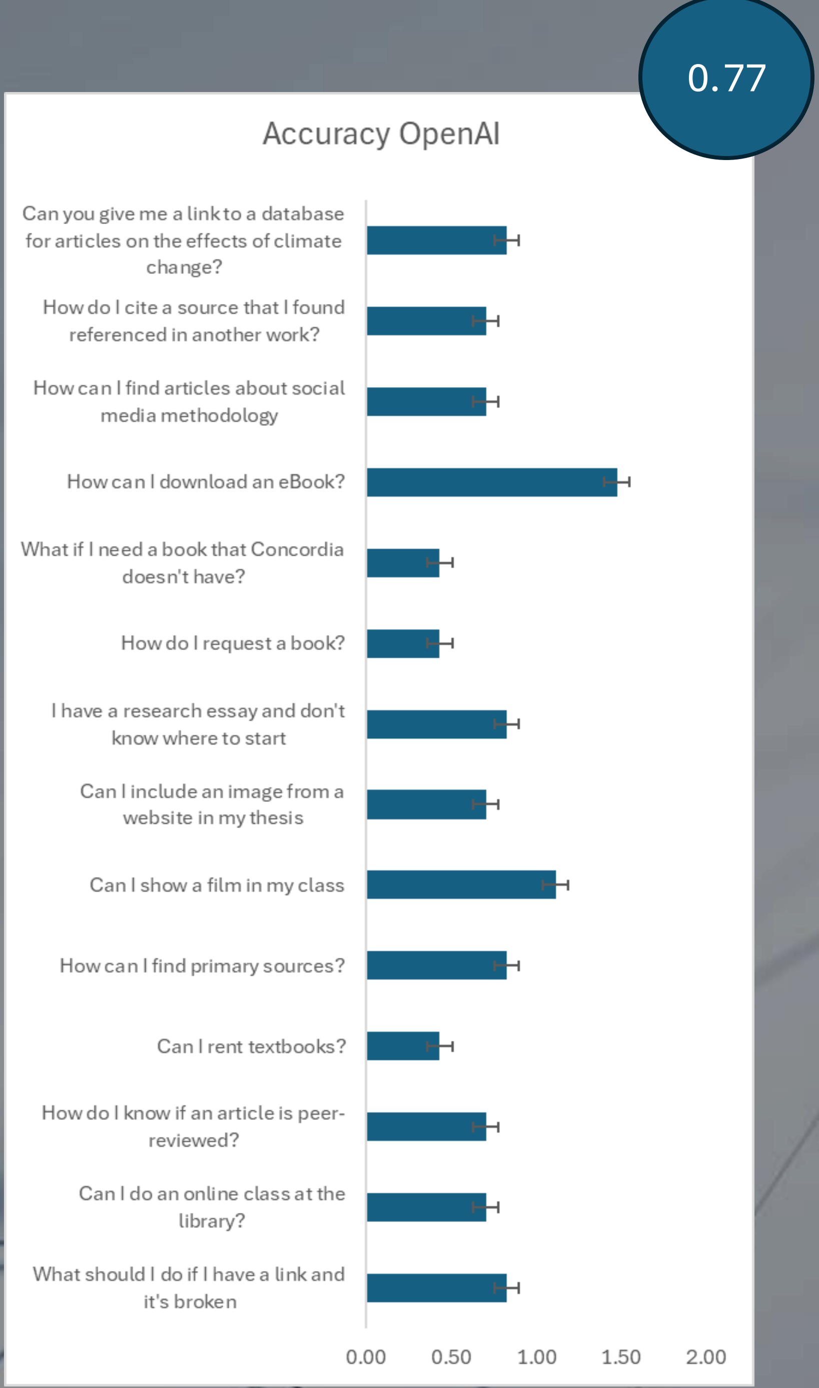
REFERENCES

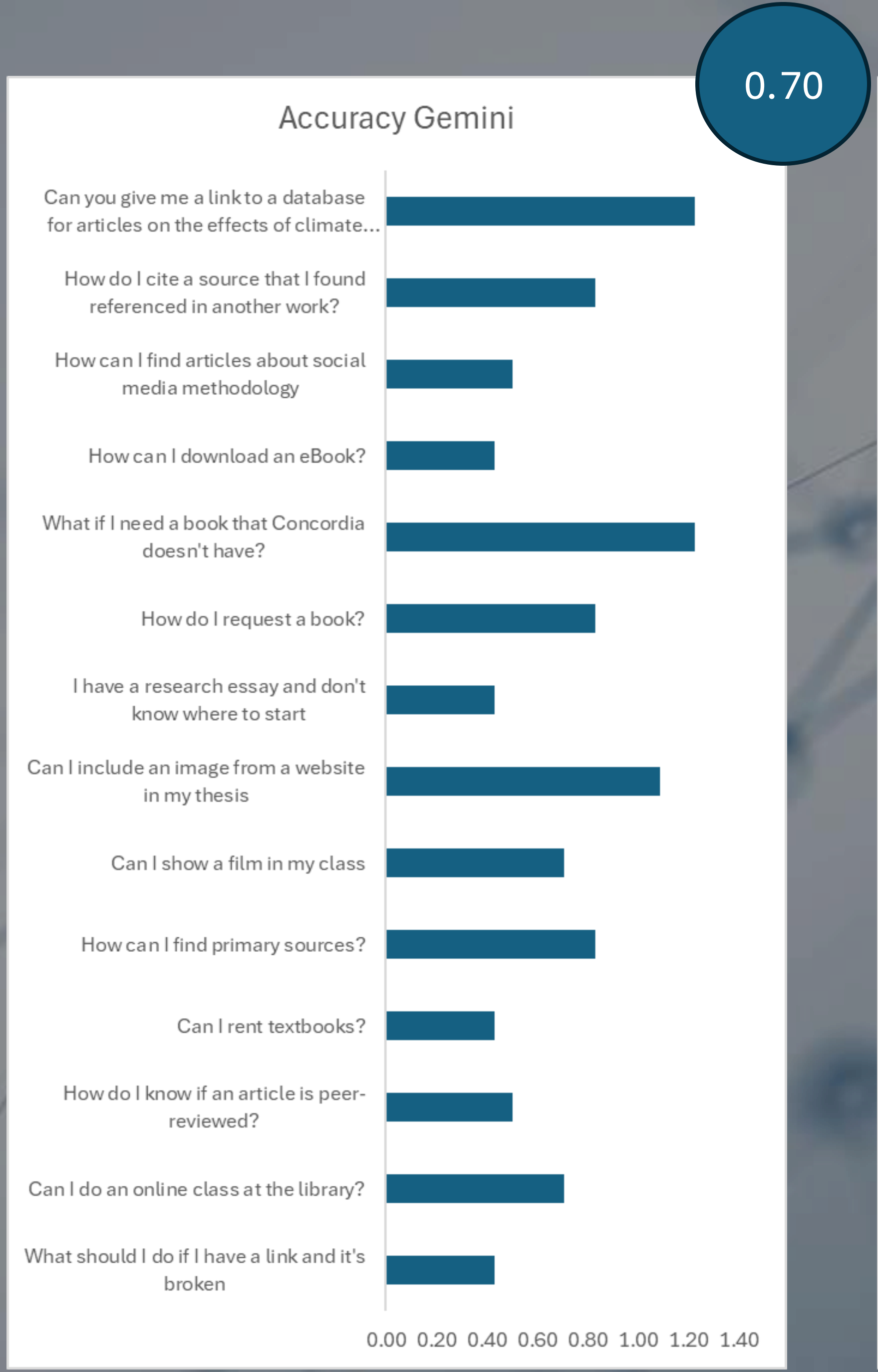
Lai, K. (2023). How well does ChatGPT handle reference inquiries? An analysis based on question types and question complexities. *College & Research Libraries*, *84*(6), 974-995. https://doi.org/10.5860/crl.84.6.974

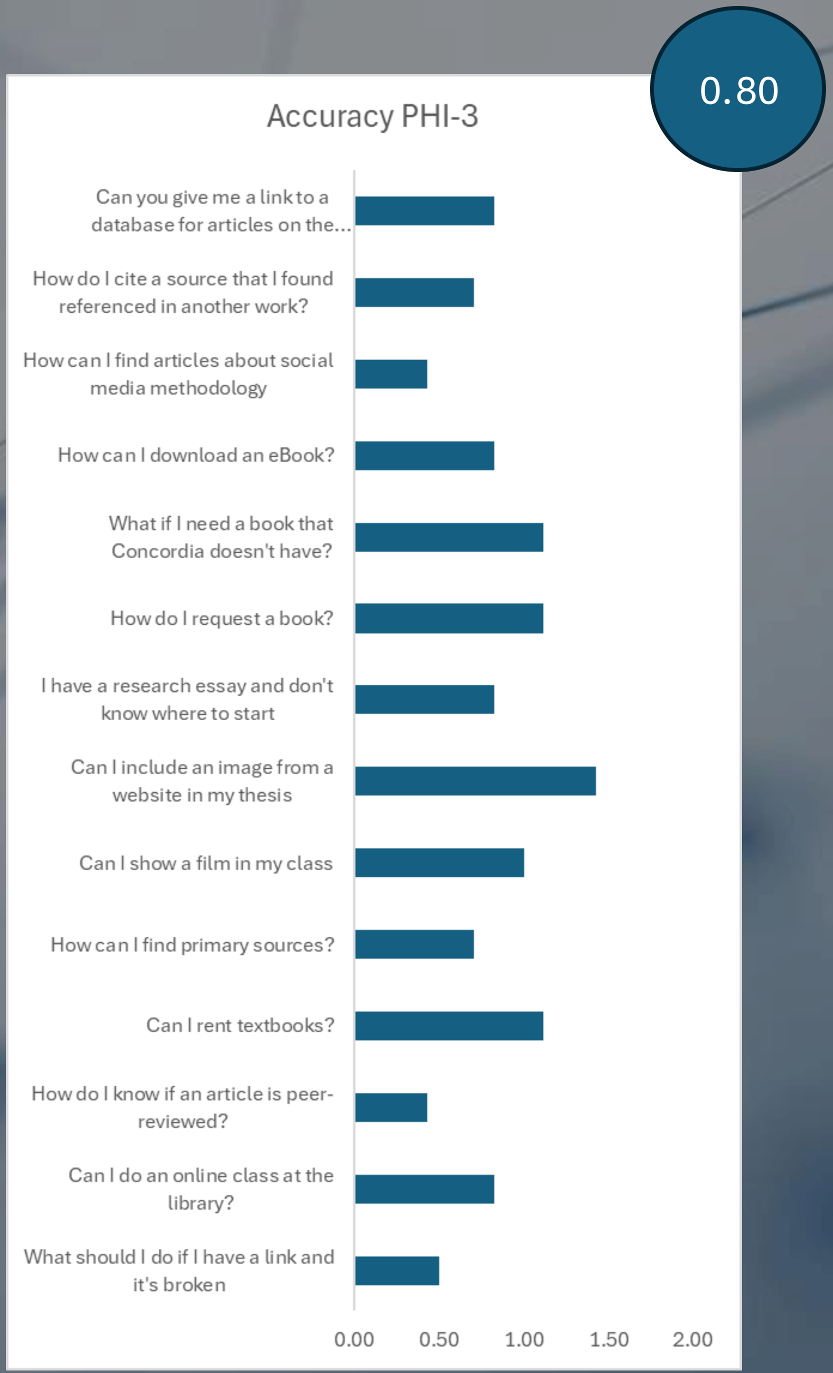
Lappalainen, Y. & Narayanan, N. (2023). Aisha: A custom AI library chatbot using the ChatGPT API. *Journal of Web Librarianship 17(*3), 37-58. https://doi.org/10.1080/19322909.2023.2221477

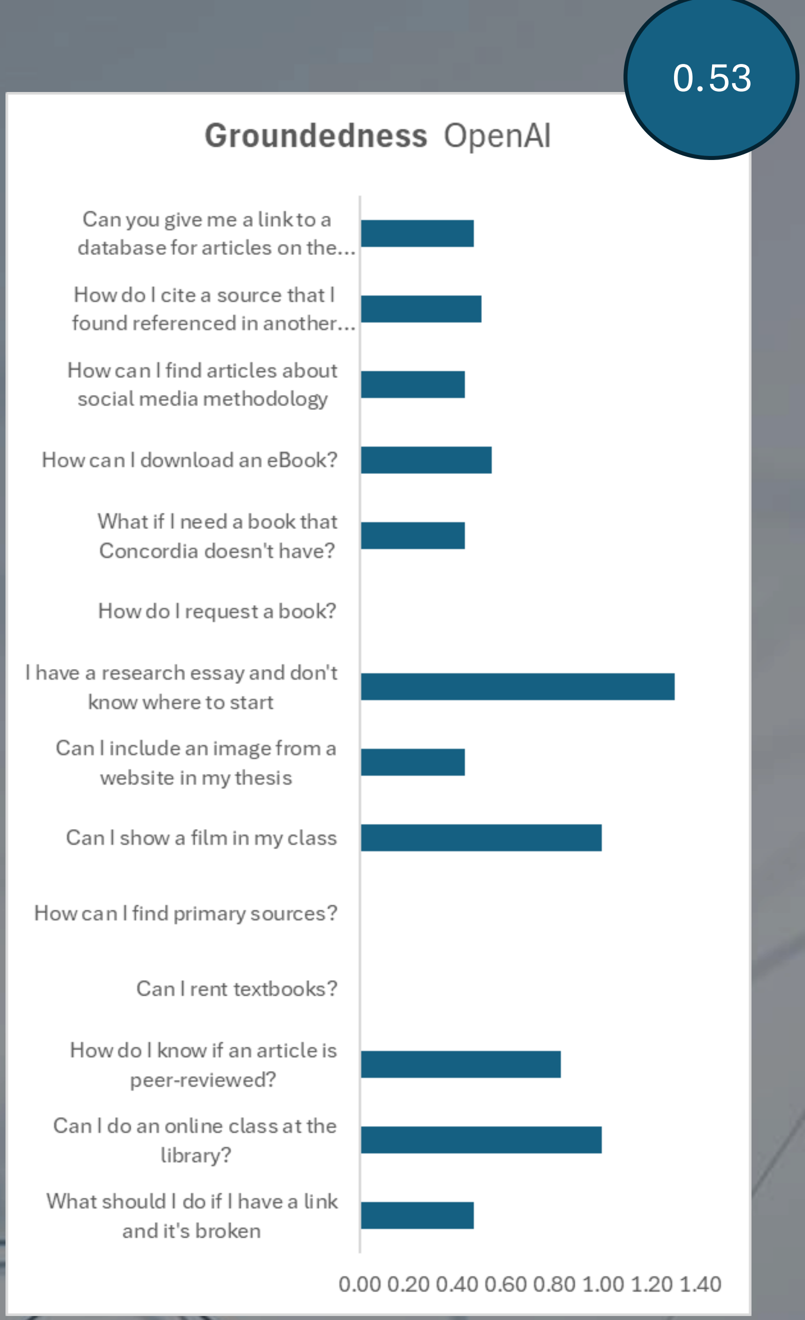
Library of Congress. (2023) LC Labs AI Planning Framework. Reprint, *Library of Congress*. https://github.com/LibraryOfCongress/labs-ai-framework.

THANKS

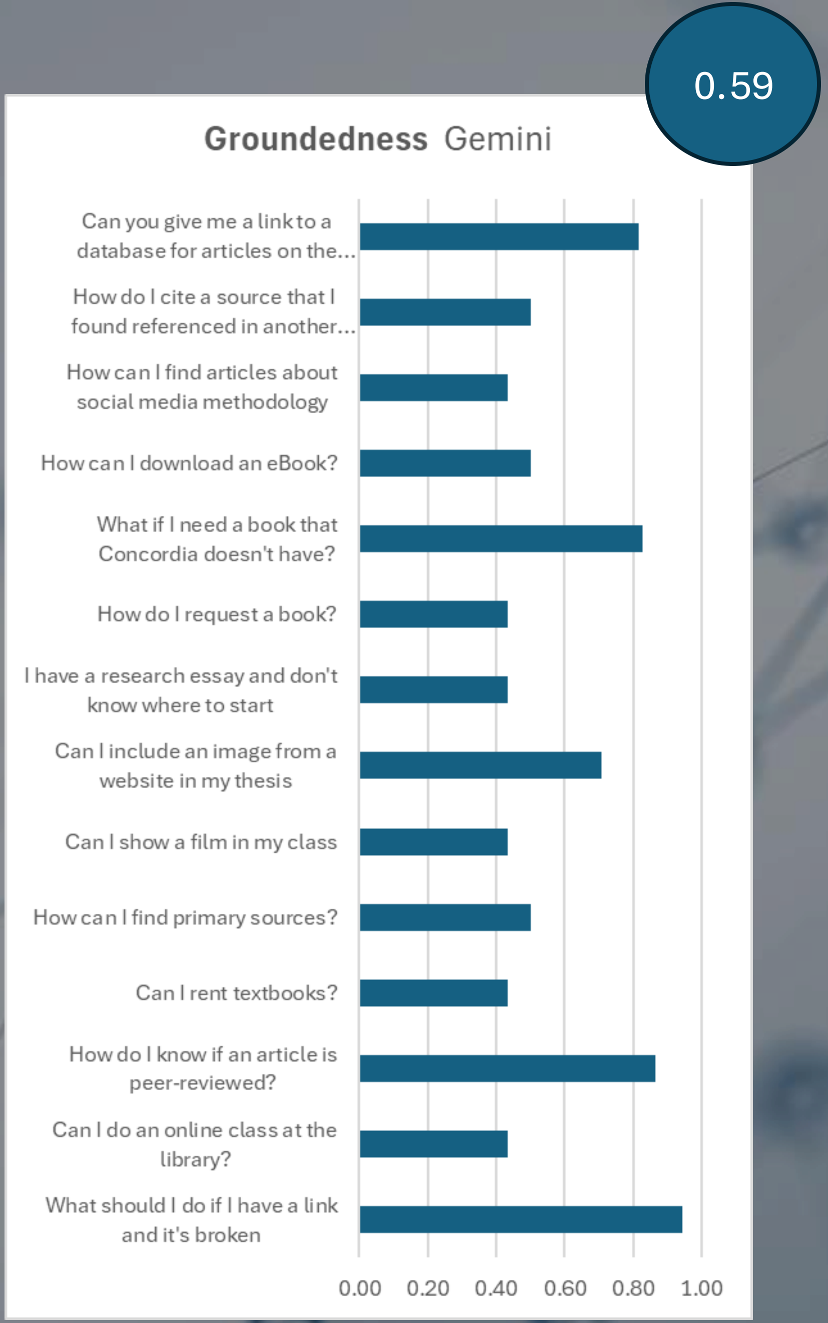

0.77
Accuracy OpenAI
Can you give me a link to a database for articles on the effects of climate change?
How do I cite a source that I found referenced in another work?
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.50
1.00
1.50
2.00


0.70
Accuracy Gemini
Can you give me a link to a database for articles on the effects of climate...
How do I cite a source that I found referenced in another work?
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40

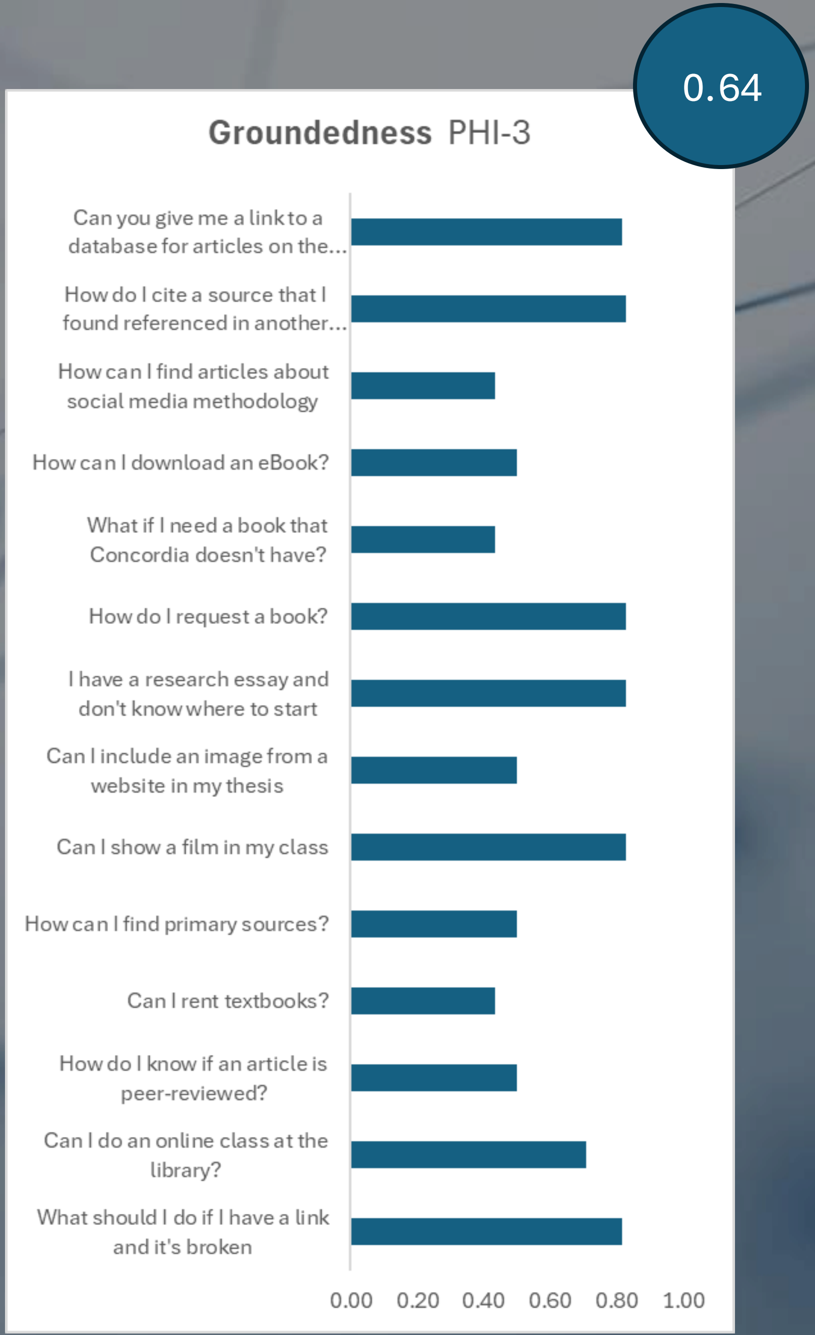

0.80
Accuracy PHI-3
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another work?
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.50
1.00
1.50
2.00


Groundedness OpenAI
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another...
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00 0.20 0.40 0.60 0.80 1.00 1.20 1.40

0.53

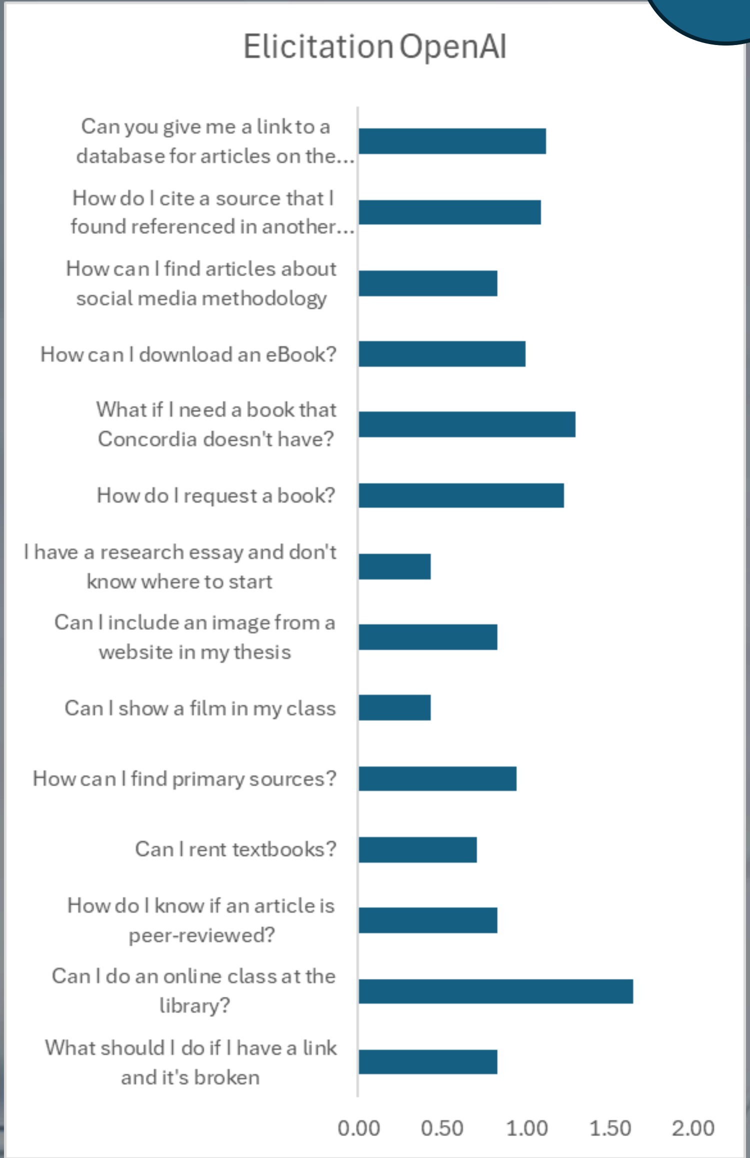

Groundedness Gemini
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another...
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.20
0.40
0.60
0.80
1.00

0.59


Groundedness PHI-3
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another...
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.20
0.40
0.60
0.80
1.00

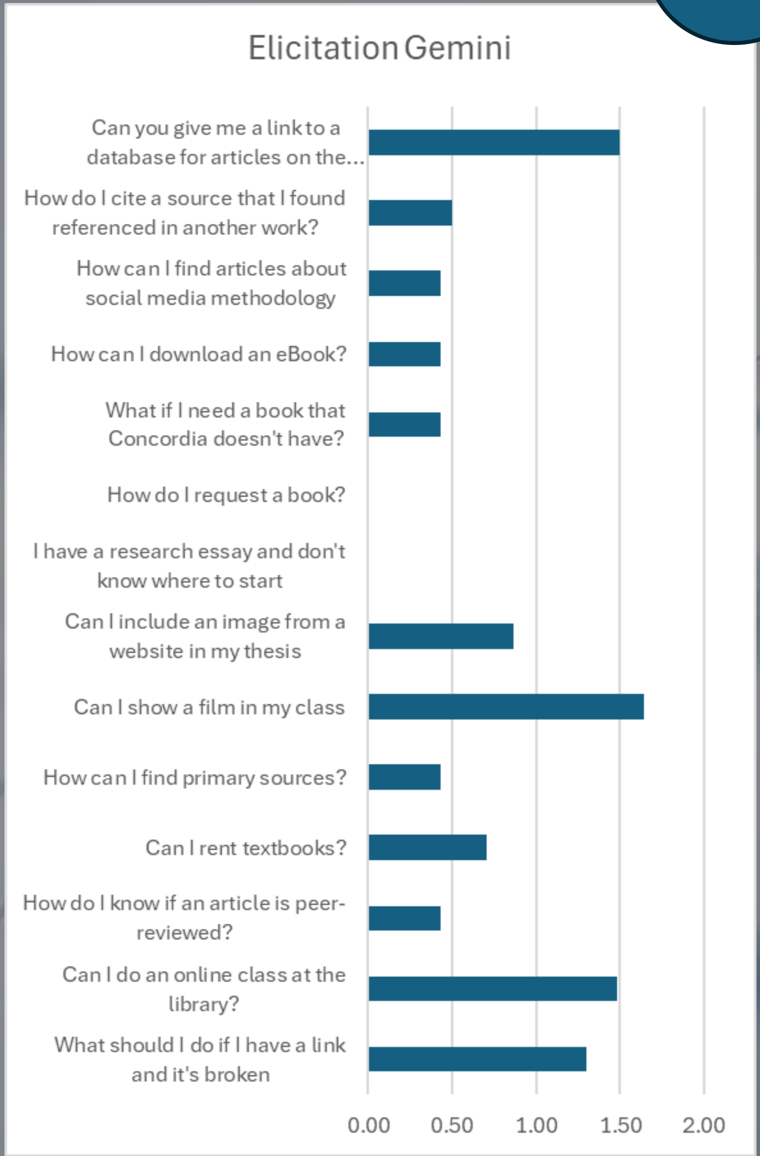
0.64

0.94



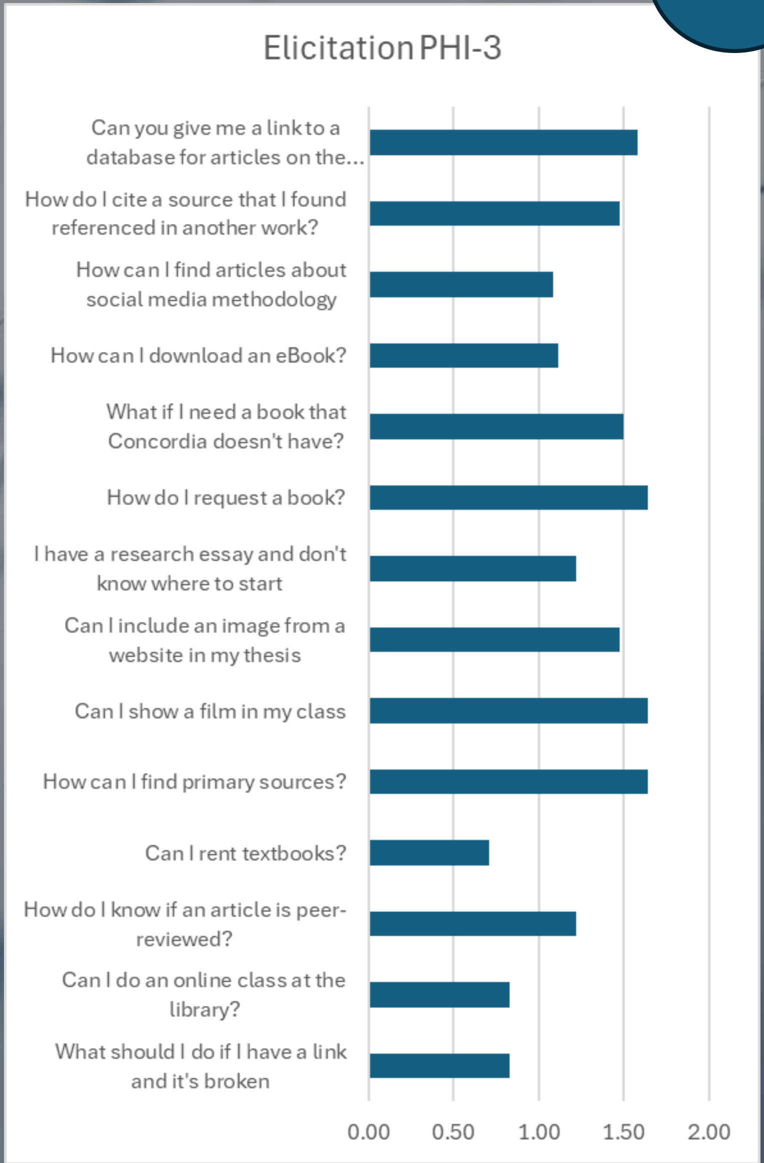
Elicitation OpenAI

0.73

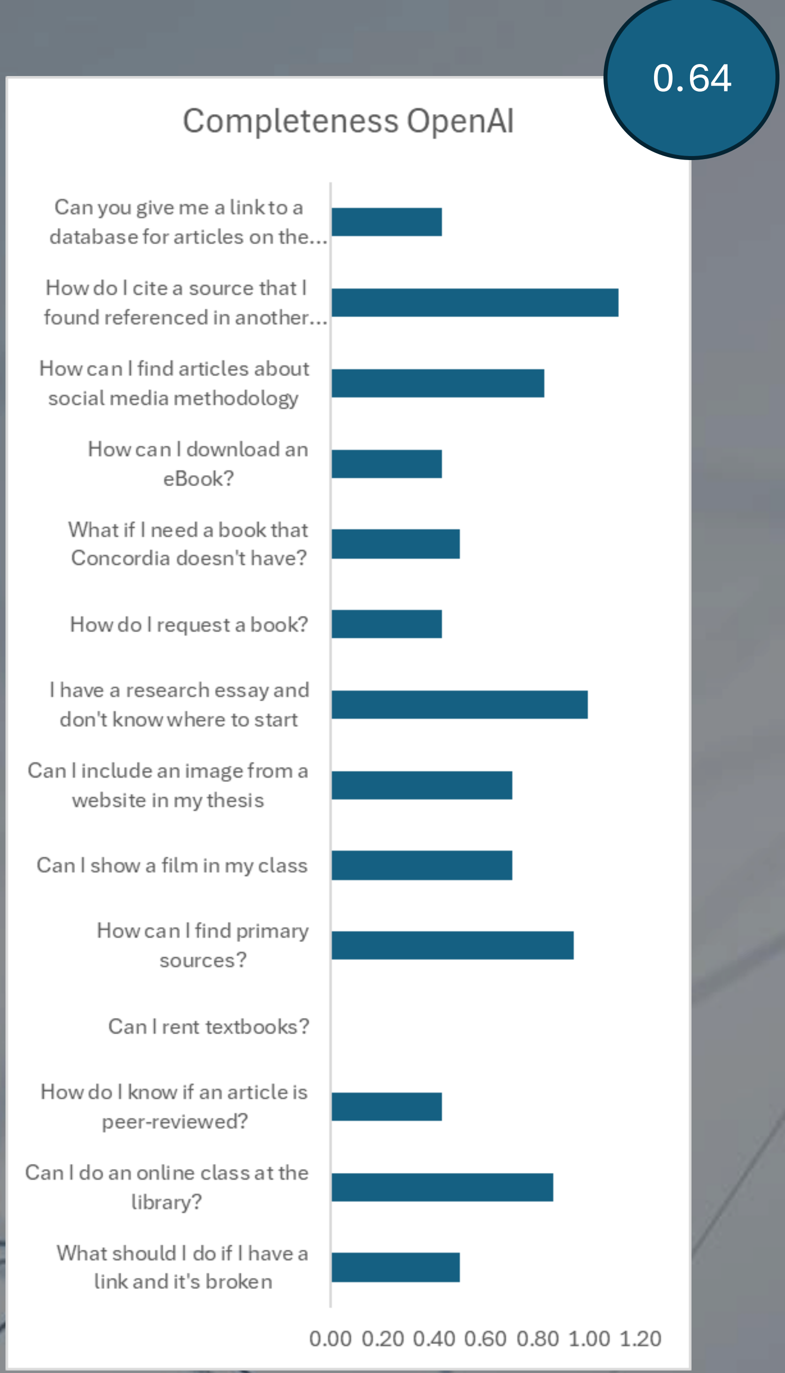


Elicitation Gemini

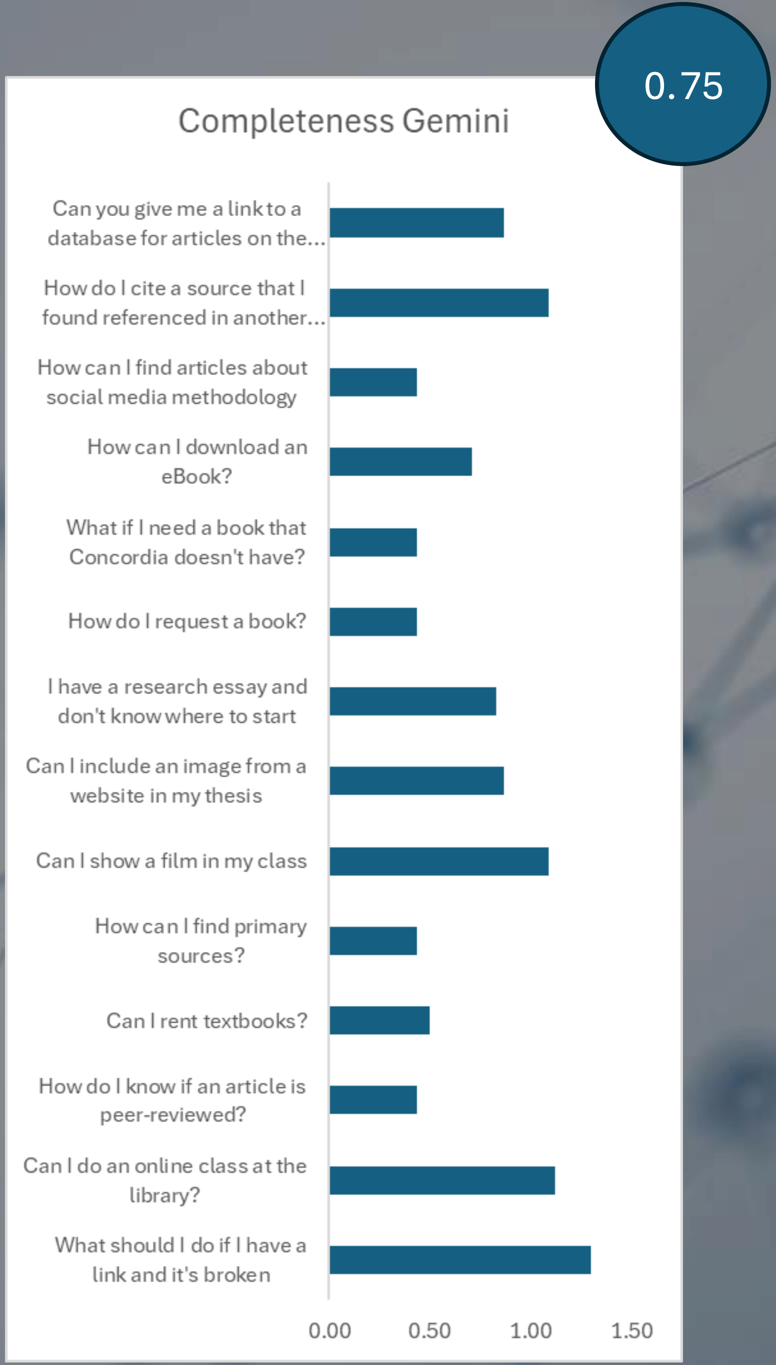
1.28



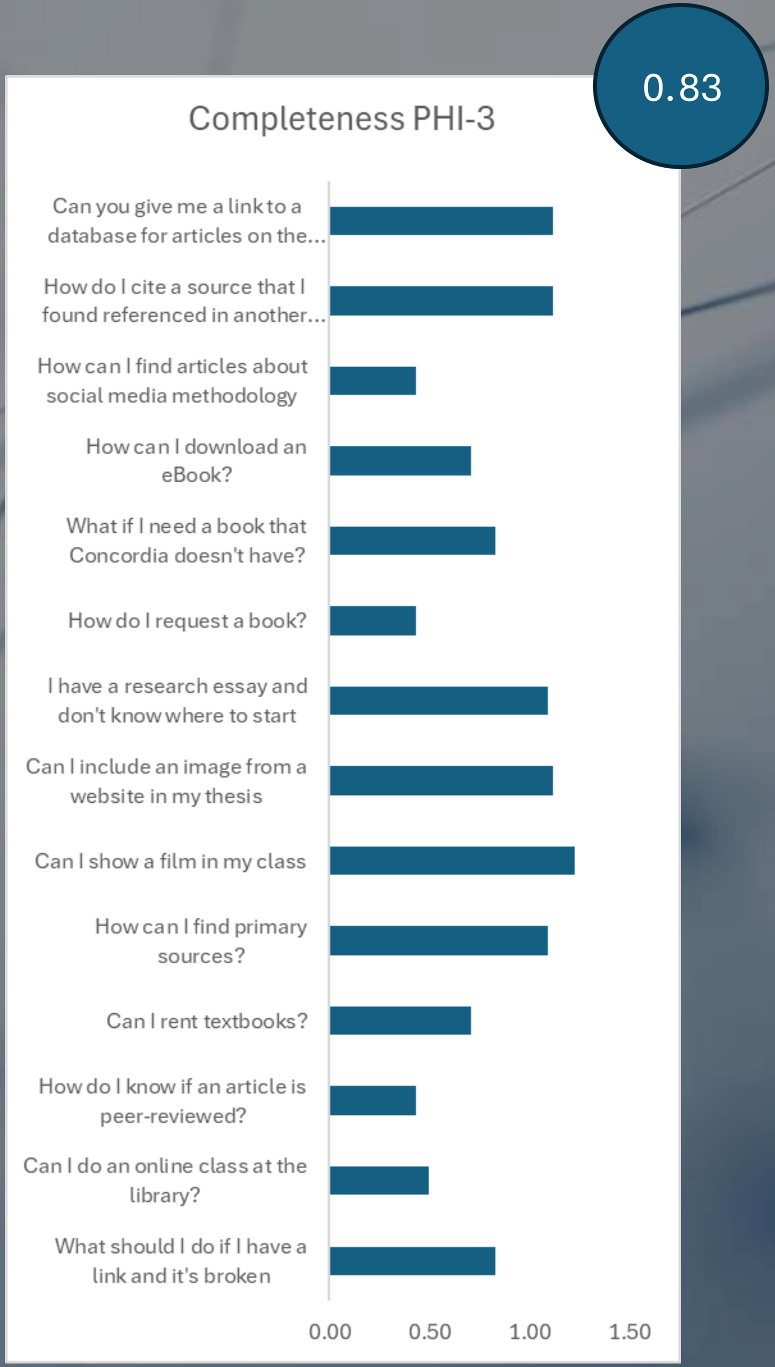
Elicitation PHI-3


Completeness OpenAI
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another...
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00 0.20 0.40 0.60 0.80 1.00 1.20

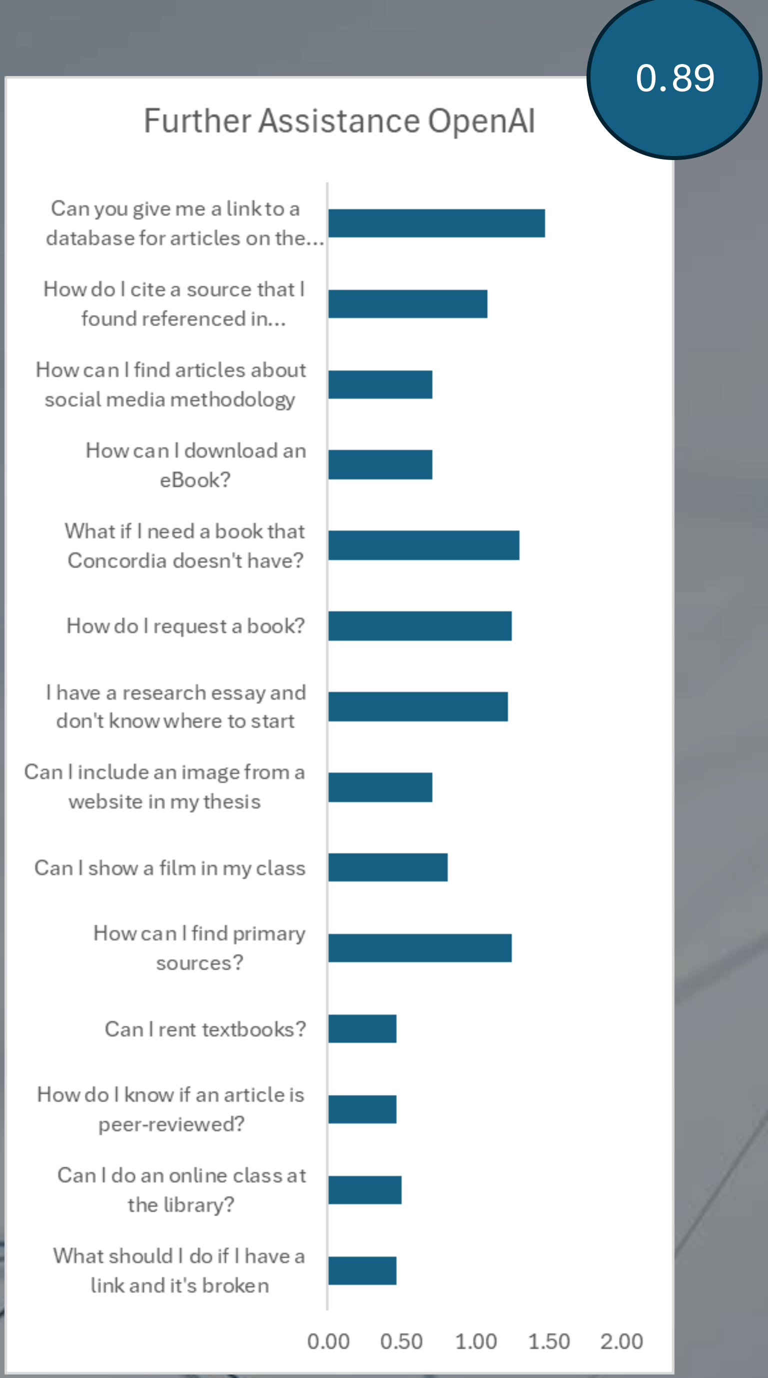
0.64


Completeness Gemini
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another...
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.50
1.00
1.50

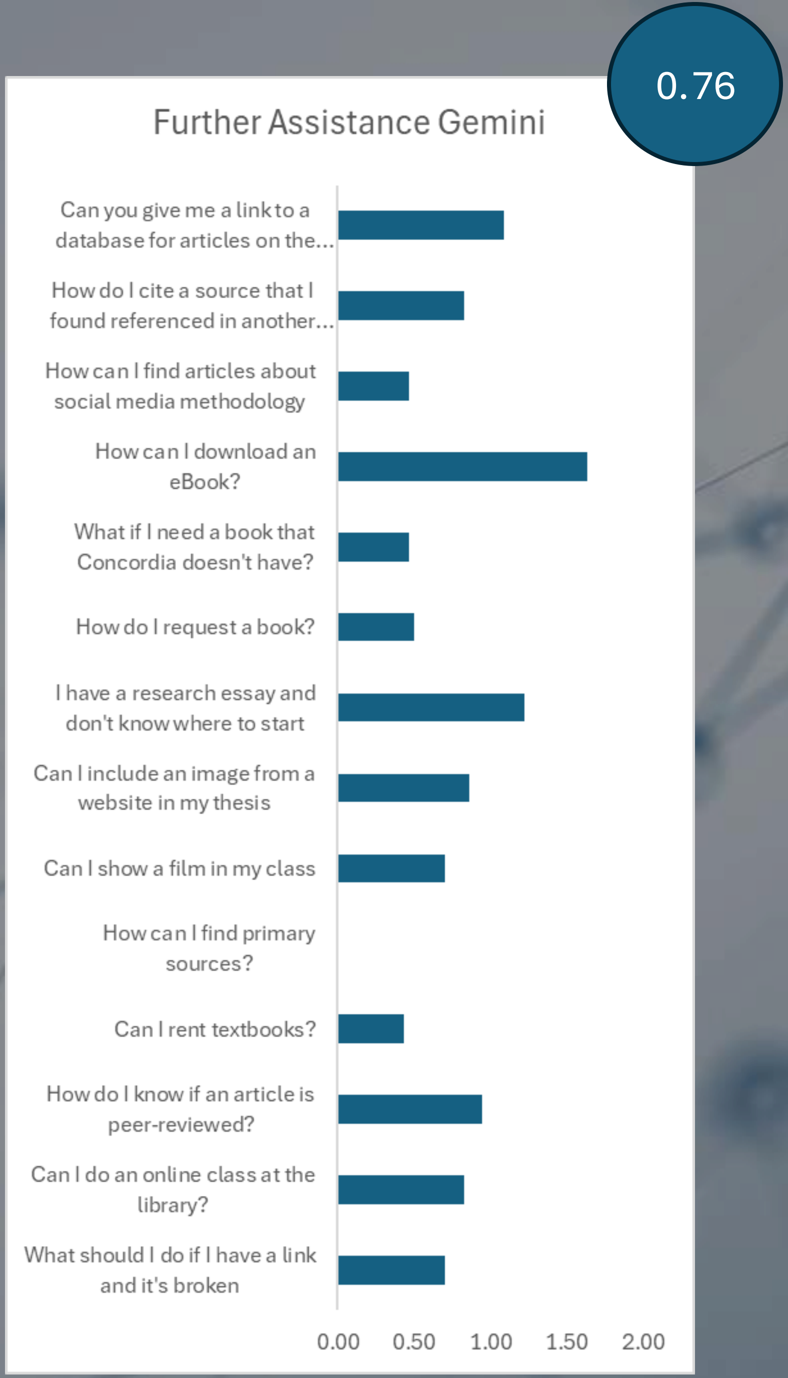
0.75


Completeness PHI-3
Can you give me a link to a database for articles on the...
How do I cite a source that I found referenced in another...
How can I find articles about social media methodology
How can I download an eBook?
What if I need a book that Concordia doesn't have?
How do I request a book?
I have a research essay and don't know where to start
Can I include an image from a website in my thesis
Can I show a film in my class
How can I find primary sources?
Can I rent textbooks?
How do I know if an article is peer-reviewed?
Can I do an online class at the library?
What should I do if I have a link and it's broken
0.00
0.50
1.00
1.50

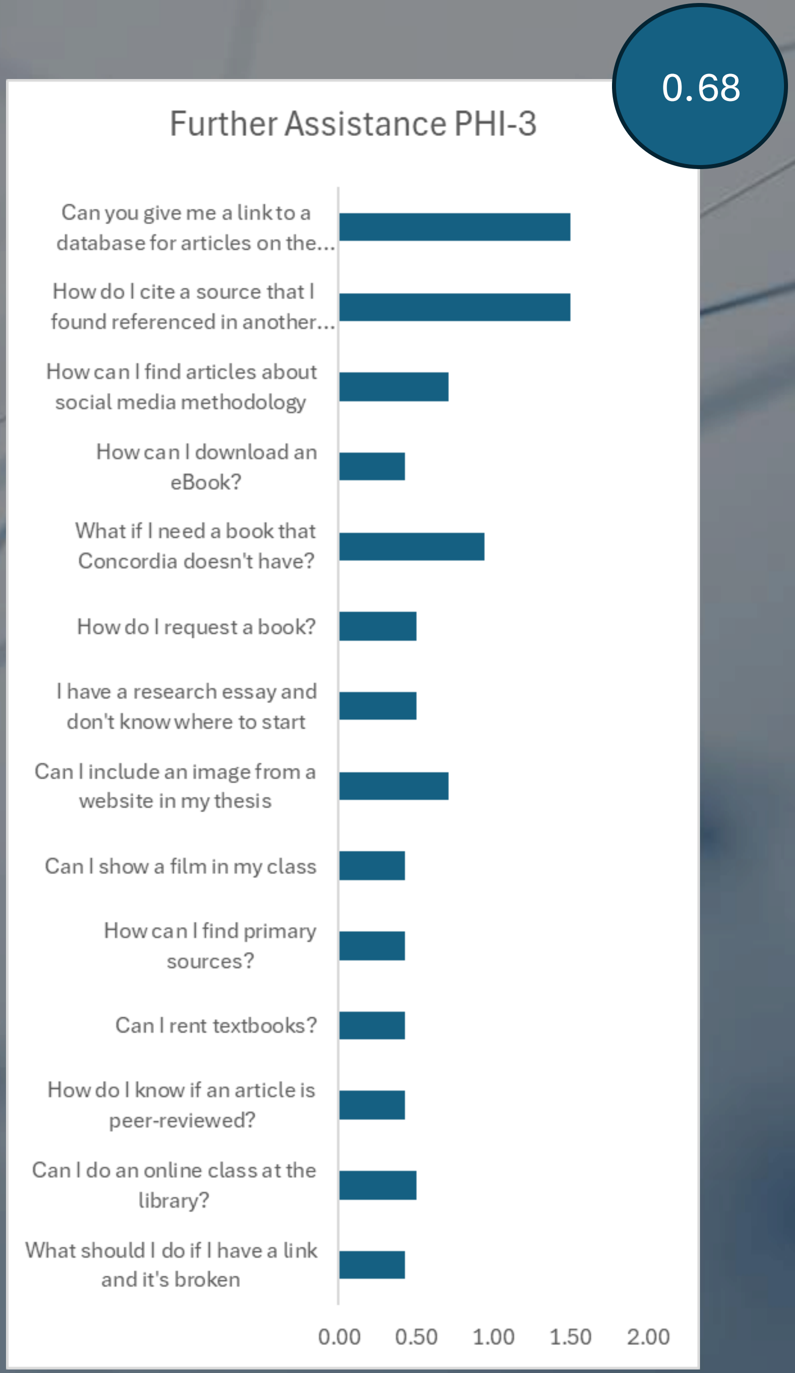
0.83


Further Assistance OpenAI

0.89


Further Assistance Gemini

0.76


Further Assistance PHI-3

0.68