

Using ChatGPT as a Tool for the Acquisition and Assessment of E-Resources

Amanda Yesilbas

Abstract

The use of artificial intelligence products like ChatGPT is not yet common in librarianship. Librarians may be hesitant to adopt such tools due to concerns that the new technology might threaten their perceived usefulness as professionals and ethical issues surrounding the use of generative artificial intelligence. This proceeding offers a historical perspective on how technologies, once viewed as threats, eventually became essential for libraries. It serves as an introduction to ChatGPT for those unfamiliar with the product, demonstrating some potential uses in acquisitions and collection assessment through various examples. The author acknowledges the challenges and potential pitfalls of using ChatGPT, highlighting the difficulties of integrating new and unproven technology. Despite these challenges, artificial intelligence tools like ChatGPT are available, and librarians are encouraged to become informed users and explore how they can enhance workflows.

Keywords: ChatGPT, artificial intelligence, e-resources, acquisition workflows, collection assessment

The Pew Research Center found in February 2024 that on average only 23 percent of American adults have used ChatGPT.¹ By extension this likely means most American librarians have not used ChatGPT or applied it to their work. Part of the reluctance to explore the uses of ChatGPT or other commercially available artificial intelligences (AI) might be due to the existential dread surrounding these products. It is difficult to embrace a product that could purportedly replace the work of librarian and information professionals, jeopardizing their careers, and especially a product that is plagued by ethical concerns about the training data and the accuracy of the output.

The first part of the presentation will provide a brief history of technologies that were initially thought to threaten libraries, but ultimately became integrated tools. It theorizes that AI tools will follow a similar trend. The second part of the presentation will introduce and walk through the product ChatGPT for audience members who have not experienced it before, leading to a demonstration of how ChatGPT could be used to assist with acquisition and assessment work in a library setting.

Timeline of Technology and the Anticipated Death of the Library

The fear of new technology is a common theme in the history of librarianship. Each new technology has significantly changed the way librarians work, but the profession has nonetheless willingly integrated these once-threatening tools. The process of automating libraries, for example, was initially met with fear. In 1964, the *Wilson Library Bulletin* published a symposium titled "Intelligent Woman's Guide to Automation in the Library" to explain the value of automation in layperson terms. As the editor Jesse H. Shara assuagingly convinced readers: "automation does not exist to give machines something to do or to put people out of work."² If we move beyond the inherent sexism that the fearful were uninformed women, the symposium sheds light on the various threats that the library profession perceived at that time.

During the late 1990s and early 2000s, personal computing and the early Internet were perceived as potential threats to the existence of libraries. There were conflicting predictions about the future of libraries during that time. Some believed that libraries would become obsolete because everything was freely available on the Internet, while others recognized an ongoing need for libraries in their provision of access to computers and the Internet, as well as offering training to their communities. It was also a Wild West for the early search engines. Prior to the dominance of Google, various search engines such as Alta Vista, Dogpile, and Yahoo vied for users by searching different indexes and returning different results. Ultimately, Google won the great search engine war. The current AI market is in a comparable stage, where major technology players are striving to develop their own AI tools trained on different sets of data, leading to different results. As of now, there is no clear indication of which AI product might dominate as development continues.

The idea that the early Internet might replace libraries seems improbable in hindsight. During this time databases existed on CDs, websites did not have search features, primitive online journals came as single access add-ons to print subscriptions, and most importantly content was not free on the Internet. Libraries, as always, were still paying and providing access.

With the evolution of the Internet, platforms like Wikipedia, YouTube, social media, and streaming services have become popular. Many people might have thought that this would mark the end of libraries. However, libraries have arguably grown even more crucial in the modern era. They offer reliable and curated information, promote information literacy, and continue the traditional practice of making information easily accessible. Additionally, libraries continue to invest in diverse formats to cater to the needs of their users.

The rise of Open Access (OA) publishing is seen as a potential threat to traditional libraries. The concept suggests that if high-quality, peer-reviewed content is freely available to everyone, libraries may become obsolete. This overlooks the nuances of library

agreements with publishers, however, the number of libraries that continue to subscribe to hybrid journals and the degree to which libraries are engaged as publishers. Furthermore, it fails to acknowledge the many other valuable services that libraries offer to their communities.

Using ChatGPT for Acquisition and Assessment

The author offered this exploration of technology in the library to set up the argument that instead of a potential threat, AI should be viewed as a potential tool to be used in library work. In this part of the presentation the author will demonstrate how ChatGPT might be used in acquisition and assessment workflows.

Vocabulary

Before discussing ChatGPT, the author will first define some vocabulary.

Large language models (LLMs) are deep learning algorithms that can recognize, summarize, translate, predict, and generate content using very large datasets. Large language models largely represent a class of deep learning architectures called transformer networks. A transformer model is a neural network that learns context and meaning by tracking relationships in sequential data, like the words in this sentence.³

Generative AI is a type of artificial intelligence that can create new content, such as text, images, music, or even video, based on patterns and data it has learned. Instead of just analyzing and understanding existing data, generative AI produces original outputs, often mimicking human creativity.

A prompt is a natural language request for information from a generative AI.

GPTs are customizable versions of ChatGPT made for specific tasks.

What ChatGPT Looks Like

It is difficult to envision how a product works if you have never seen it. Figure 1 shows what ChatGPT looks like when the webpage is opened at the time of writing. ChatGPT is a rapidly developing product that undergoes regular changes. Each number highlights a point of interest.

1. This is where the prompt is entered. The paper clip indicates where an attachment can be uploaded. I will use this feature to upload spreadsheets into ChatGPT for the following examples.
2. This indicates what version of ChatGPT is in use. It can be changed to earlier versions.
3. This indicates which GPT is being used. GPTs are chats that have been customized to do specific things like create images or analyze data. The examples in this presentation use the default GPT as some GPTs like Dalle are only available on paid accounts.
4. This shows a history of previous chats. Previous chats can be reopened and ChatGPT now supposedly remembers previous conversations.



Figure 1. ChatGPT Screenshot

In the presentation, the video, https://youtu.be/uaX4R0CH_Yc, is played to demonstrate ChatGPT in action, allowing viewers to see the product being used.

Tips for Working with ChatGPT

ChatGPT learns as it works with the user. A useful analogy for using ChatGPT is to treat it like a new employee. In the beginning of a chat, use explicit instructions to prevent confusion. If a new spreadsheet is loaded into the chat, ask ChatGPT to use data from the exact column name. For example, if the new spreadsheet has a column called "Total Cost," ask ChatGPT to work on the "Total Cost" column for the first interactions. After the chat has trained on your data, it now understands that the data is cost data. Later in the conversation, prompts like "What is the cost by vendor?" will return the correct information without asking specifically for "Total Cost." The chat will learn the data once it is defined and can be interacted with in a more casual manner with less explicit instruction.

It is crucial for the user to understand the data that ChatGPT is working with. Like a new employee, ChatGPT can initially make mistakes if it does not understand the data or the questions asked. Sometimes it can be incorrect without a clear reason. The user should be able to review the results provided, ensure they make sense, and make any necessary corrections to the prompts.

Once ChatGPT comprehends the data and the user's intent, it can function more like a trusted colleague. Users can ask it to perform complex tasks without providing detailed instructions. It can be interacted with in a conversational manner, assisting the user in problem-solving and carrying out intricate calculations without the need to open a spreadsheet.

After all the data analysis is done the chat can then perform managerial tasks like writing executive summaries and reports.

Library Use Cases for Using ChatGPT

The next portion of the presentation walked audience members through how to do specific tasks primarily related to acquisitions and assessment work.

Anonymizing Data

In this example, a spreadsheet with e-book purchases for a five-month period was uploaded into a chat. The author needed to anonymize the vendors, but was not sure how to do it using formulas. Instead, the author explained to ChatGPT what she wanted to do. She began by telling the chat that the “Publisher” column was a list of publisher’s names. To anonymize the names the author wrote the prompt, “I need to anonymize the names but the names must remain consistent. Example Guilford Publications must be the same anonymized name in each instance.” The chat returned properly anonymized names with Guilford Publications being renamed to Publisher_3 in all instances.

Randomizing Data

In the same chat session where the vendor’s identity was anonymized, the author aimed to demonstrate how to easily make mass changes to a spreadsheet by randomizing the prices. To simplify the task for both the user and the chat, the author divided the process into two separate prompts. The first prompt instructed, “I need to randomize the Unit Price. If the value is one-hundred or over, randomly add or subtract an amount between twenty and fifty.” After successful completion of this step, the author issued the second prompt: “If the unit price is less than one-hundred, randomly add or subtract an amount between five and ten. However, the value should not go below one.” Despite grammatical errors, the chat properly executed the request. It

is worth noting that ChatGPT is very forgiving of spelling errors, typos, and bad grammar.

Monthly Expenditures

In the same chat, a request was made for spending by month. ChatGPT returned a table of the amount spent by month, demonstrating how data can be easily pulled and manipulated using simple commands, without having to open a spreadsheet.

Expenditure by Subject

Creating an expenditure report by subject proved to be challenging due to the way the data was entered into the original spreadsheet. This example emphasizes why librarians need to be knowledgeable about the data input and vigilant in analyzing ChatGPT's results.

The prompt "What subjects had the highest spend in 2023?" returned a seemingly correct table, but the numbers seemed too low. Looking at the original data entered into the chat showed that the subject column contained multiple subjects in a single field. As a result, ChatGPT was only totaling columns that had a single value. For example, Social Science was the highest spend, but the chat only included e-books that had the single subject of Social Science. It did not include titles with multiple subjects like "Social Science; History" or "Social Science; Literature." The totals were incorrect because not all Social Science e-books were included in the total spend.

To overcome this, the author needed to group together titles by subject. She started by just grouping Social Science titles together. The author's philosophy is to test and expand. In this instance she wanted to make sure the Social Science grouping worked before trying to group together all the other subjects. The first prompt given was "Any value in the column Subject heading that starts with the words Social Science is to be grouped into a category called Social Sciences." This returned a new table with the heading "Grouped Subject"

and another column "Amount Spent (\$)." The expanded prompt was "can you do the same for each value in the grouped subject field? If any value in the subject column starts with education group them all in a group education. If any values in the subject column start with Literature group them together in a group literature. Do this for all the values in the Grouped Subject column"

While ChatGPT did successfully complete the work and return the correct amounts (the author checked), the author now realized the prompt was flawed in design asking to group by first subjects. If the subject has appeared second in the column, it might not be properly grouped and counted. ChatGPT did as asked but the user was in error so the spend is probably not correct. ChatGPT is only as good as the operator.

Creating Visualizations

The next portion of the presentation demonstrated how ChatGPT can make graphs and other data visualizations just by asking. It does not require formulas or pivot tables.

Creating Summaries

The last example working with demonstrated how ChatGPT could easily summarize the data in a chat into an executive report or outline.

Bonus Example

In the final example, the focus was on cataloging. The author demonstrated how two different file types could be uploaded and used together. In this case, an Excel file was uploaded containing e-book titles with subject headings but no call numbers, and a Microsoft Word file with the Library of Congress Classification System outline was also uploaded. The author asked ChatGPT to look up subject headings from the Excel file using the information from the Word document and then return a spreadsheet with the call numbers added.

Failures

After demonstrating use cases, the author showed examples of failures that occurred while using ChatGPT. These included instances where no data was included in a table, and updates and saves did not work properly. Additionally, ChatGPT had trouble processing a PDF file due to its size, but it was able to accurately process the same file in Word format.

Warnings

Although ChatGPT has some promising features, it comes with many problems as well.

- It can make up information.
- It can be lazy and refuse to complete tasks without coercion.
- It is gullible and can be compelled to do things against usage terms with creatively worded prompts.
- It is biased. It only knows what it has been trained on and sadly the information fed to it contains biases it repeats.
- It has usage limits and will kick a user off if they do too much in a period time. It will also throttle users during peak times.
- Privacy is not guaranteed. Be careful of the data inputted.

In Conclusion

Why should we use ChatGPT or another AI tool despite their problems? The truth is, AI is becoming increasingly prevalent. It is already being integrated into everyday products used by librarians, from word processing to databases. Librarians should strive to become educated users of AI, and by using these tools, we might be able to shape them to better suit our needs and profession.

At this time the author is not fully comfortable using ChatGPT for professional purposes beyond mathematical calculations or routine tasks. However, even for these functions, AI can simplify and speed up repetitive tasks and workflows. Commercially available AI products are improving continuously, and it may be beneficial to see AI products as just another tool to leverage.

Contributor Notes

Amanda Yesilbas is Electronic Resources Librarian, Assistant Librarian at the University of South Florida Libraries, Tampa, FL, USA.

Notes

- 1 Colleen McClain, "Americans' Use of ChatGPT Is Ticking Up, but Few Trust Its Election Information," Pew Research Center, March 26, 2024, accessed July 31, 2024, <https://www.pewresearch.org/short-reads/2024/03/26/americans-use-of-chatgpt-is-ticking-up-but-few-trust-its-election-information/>.
- 2 Jesse Hauk Shera, ed., "Intelligent Woman's Guide to Automation in the Library, a Symposium," *Wilson Library Bulletin* 38 (May 1964): 741–779.
- 3 "Large Language Models Explained," NVIDIA, accessed May 28, 2024, <https://www.nvidia.com/en-us/glossary/large-language-models/>.