

AI for Archival Description:

Prompt Engineering for Archivists

Outline



1. **Prompt engineering:** definition and components of a well-crafted prompt for ChatGPT
2. **Examples** of successful prompts
3. **Managing privacy**
4. **Vetting AI outputs** for inclusion in archival catalogs
5. **Crediting AI** for assistance with description
6. **Resources**

Prompt Engineering



Prompt engineering is the practice of creating effective prompts that allow generative artificial intelligence (AI) tools to produce desired outputs. Prompt engineering can be useful to most professions, but for archivists specifically, learning how to effectively write prompts for AI tools like ChatGPT can automate otherwise time-consuming tasks. This poster centers on prompt engineering for **ChatGPT** to help archivists quickly and accurately create **scope and content notes**.

Some **components of effective prompts** include:

- **Expressly state any rules that the desired output must follow** e.g., tell ChatGPT whether you want the response to spell out acronyms, comply with *Describing Archives: A Content Standard (DACS)*, and incorporate elements of accessible description.
- **Provide all necessary resources** for ChatGPT to produce a desired output without need for external knowledge.
- **Clearly defined specific instructions:** state the obvious, such as clarifying that you do not want ChatGPT to make anything up. Also be sure to provide a sufficient explanation of the task you want the AI to perform.
- **Example scenarios (optional):** examples will teach ChatGPT what kind of output you want. I find that this component, while helpful, is not always necessary to produce a high-quality output.

ChatGPT 3.5 (free version): Examples of successful prompts



Characteristics of ChatGPT 3.5:

- Can generate mid- to high-quality scope notes.
- Has limited contextual understanding, meaning you will need to provide **any/all** contextual information about the item(s) being described (e.g., copy/paste the full text of a document into the prompt).

Scenario:

You have digital or digitized non-sensitive materials in need of description. Because the materials are not sensitive, text can be copied directly from these materials. Example content type could be a news article online. Please note that there is a limit on how many characters you can place in a prompt; ChatGPT will error out if too many characters are entered. In these cases, an excerpt can be input into the prompt.

Example prompt:

Using the guidelines for archival description outlined in “Describing Archives: A Content Standard” (DACS), remaining completely objective and avoiding all value statements, ensuring that the produced description is as concise as possible, using approximately only one paragraph, provide a scope and content note for the article, **[article title]** by **[author name]** and published by the **[name of publication]**. Please also spell out all acronyms utilized in the scope and content note, and do not make anything up. The text of the article is pasted below:

- **[Paste full text of article here, if possible]**

ChatGPT 4.0 (paid version): Examples of successful prompts



Characteristics of ChatGPT 4.0:

- Can generate mostly high-quality scope notes.
- Better contextual understanding than 3.5.
- Can understand and respond to more complex queries.
- Can accept digital files directly to analyze as part of the prompt.
- Can handle aggregate description well.

Scenario:

You have digital or digitized non-sensitive materials in need of description. Because the materials are not sensitive, text can be copied directly from these materials. However, because ChatGPT 4.0 has more advanced capabilities, you can also attach documents to your prompt in lieu of copy/pasting full text of an item. This version also allows for much more text in the prompt, eliminating most text limitations of ChatGPT 3.5. Additionally, because 4.0 can accept and interpret much more information than 3.5, 4.0 is a great option for assistance with aggregate description. The following example prompt demonstrates how to effectively go about asking 4.0 to assist with description of two or more sources.

Example prompt:

Using the guidelines for archival description outlined in “Describing Archives: A Content Standard” (DACS), remaining completely objective and avoiding all value statements, ensuring that the produced description is as concise as possible using approximately only one paragraph, provide a scope and content note for a set of letters from **[sender of letters]** to **[recipient of letters]**. Please also spell out all acronyms utilized in the scope and content note, and do not make anything up. The letter transcript is pasted below/attached:

- **[Paste full text of letters here] -OR-**
- **[Upload relevant document(s)]**

Managing privacy



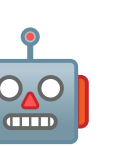
According to ChatGPT’s documentation, data and content input into the individual consumer service is used to train and improve models. Take notice of how this aligns with your institution’s data privacy policies. Enterprise/business services have different data security and ownership parameters.

Vetting outputs



Not all outputs of ChatGPT will be usable as scope notes in archival catalogs. Before adding content generated by ChatGPT to your catalog, **be sure to review for accuracy**. In some cases, the scope note might not be satisfactory, but could be a good place to start when creating your own scope note.

Crediting AI for assisted description



For transparency, it is essential to credit the AI for writing scope and content notes.

- If the archivist must edit the generated response, the following (or similar) would suffice: “Scope note composed by ChatGPT with edits by archivist.”
- If no edits have been made to ChatGPT’s output, the following (or similar) would suffice: “Scope note composed by ChatGPT.”

Resources



- DACS: <https://saa-ts-dacs.github.io/>
- ChatGPT documentation: <https://beta.openai.com/docs/>
- ChatGPT data privacy: <https://help.openai.com/en/articles/7039943-data-usage-for-consumer-services-faq>

Corey Cherrington

HT Talent | Pip512g@gmail.com

