Powered by AI:
Crafting a GPT Prototype for Oral History
Tinker with the technology

Bits and bytes align,
Crafting thoughts in circuits' dance,
New insights arise.

–GPT-4o
Creating GPT’s

Lines of code converge,
A digital mind is born,
Learning, it evolves.
–GPT-4o

Source: https://openai.com/index/introducing-the-gpt-store/
Why create a custom GPT?

- Ability to add domain knowledge
- Can focus on a specific need or topic
- Requires no coding experience
- Anyone can view / interact with custom GPT’s if set to public mode
GPT prototype: Background info

- Interview project: The “Big Top” Show Goes On
- Researchers: Tanya Finchum, Juliana Nykolaiszyn
- Repository: OSU’s Oklahoma Oral History Research Program
- Interview transcripts, audio available online since 2012
- Transcripts created, edited by humans
GPT prototype: What you need

- ChatGPT Plus account (subscription-based)
- Interview transcripts or other documents
  - File types supported:
    - Excel (.xls / .xlsx)
    - Comma-separated values (.csv)
    - PDF (.pdf)
    - JSON
Oklahoma Circus History

By Juliana M Nykolaiszon.

Based on The “Big Top” Show Goes On Oral History Project.

Tell me about a famous circus in Oklahoma.
Who are some notable figures in Oklahoma’s circus...
What major events shaped the circus culture in...
Share an interesting story from Oklahoma’s circus...

Message Oklahoma Circus History

ChatGPT can make mistakes. Check important info.
Prompts / areas for exploration

● What insights can be gained from these interviews?
● Can you create a timeline of [circus history in Oklahoma]?
● Which Oklahoma circus [still performs with elephants]?
● What are some memorable moments from [B.K. Silverlake's] interview?
● Can I ask you questions like you are [B.K. Silverlake]?
● Can you give me a list of all interviewers, interviewees, interview dates, and interview locations?
● What are some of the most asked questions?
● Could you share some key quotes from the interviews?
● What are some key topics covered in these interviews?
● You can also ask for sentiment analysis, word clouds, and more.
What about hallucinations, etc.?

- Direct quotes **from the interviews** are indeed direct quotes
- Interview locations are sometimes misrepresented
- Statistical data can be misrepresented (such as number of words per interview, etc.)
- The custom engine tends to default to a few interviews (or a few pages of transcript), instead of the entire body (hit or miss)
Thinking about ethics

- Avoid uploading files or inputting text that is not for public, online consumption
- Rethink release, consent forms
- Have discussions with your narrators about potential future use
- Read the terms and conditions for AI tools
- This is an exciting time, but cautiousness is still important
Lessons learned

- Custom GPT’s still have the ChatGPT knowledge base.
- The custom GPT may limit how many files it uses in analysis, unless specified with prompting.
- Asking questions of the entire interview set may slow down processing.
- Developing the right prompts for best results involves trial and error.
- If you notice an error, you can prompt ChatGPT to revisit.
- Always remember “ChatGPT can make mistakes. Check important info.”
What else is possible with processing?

Echoes from the past,
ChatGPT captures their soul,
Timeless words remain.

–GPT-4o
Potential uses of AI and oral history processing

- Audio transcription
- Interview summaries
- Metadata (such as Dublin Core)
- Subject headings (such as FAST)
- Brief biographies
- Cataloging
For more…

- Custom GPTs at MIT Sloan: A Comprehensive Guide
- The Reliability and Usability of ChatGPT for Library Metadata (Bodenhamer, 2023)
- An Experiment with the Use of ChatGPT for LCSH Subject Assignment on Electronic Theses and Dissertations (Chow, Kao, and Li, 2024)