Harnessing Generative AI to Support Exploration and Discovery in Library and Archival Collections

A Case Study to Incorporate Human-Centered Approaches for AI-Ready Data

**Project Team**

**RQ1:** Richard Marciano (PI): Professor
Rajesh Kumar Gnanasekaran: Doctoral Candidate, iSchool, U. of Maryland, campus IT director

**RQ2:** Mark Conrad: Digital Archivist, AI-Collaboratory (formerly with the U.S. National Archives)

**RQ3:** Lori Perine: AIC Research Fellow & former Associate Researcher in Trustworthy AI at the National Institute of Standards and Technology (NIST).
Project Overview

- Proposal to the IMLS National Leadership Grant for Libraries (NLG-L) program (April 2024)

- NLG-L Goal 3 & Objective 3.2: ‘Improve the ability of libraries and archives to provide broad access to and use of information and collections’ & ‘Support innovative approaches to digital collection management.’

- Community Partners/Collaborators:
  - Maryland State Archives
  - Legacy of Slavery Program
  - Kennard African American Cultural Heritage Center
    (Queen Anne’s County, Maryland)

Research Methodology: Exploratory Case-Study with Stakeholder Participation

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Research Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RQ1) How should we further curate library and archival collections to make them AI-ready?</td>
<td>Research collection curation techniques and AI-readiness to support generative AI work and refine GPT models to understand a collection’s context.</td>
</tr>
<tr>
<td>(RQ2) How can we compare GPT models with traditional, non-AI exploratory data analysis models?</td>
<td>Interrogate the DTA dataset using traditional, non-AI approaches for comparison with GPT model results.</td>
</tr>
<tr>
<td>(RQ3) How can we incorporate socio-technical considerations to promote trustworthiness and mitigate potential bias arising from the use of GPT models with library and archival collections?</td>
<td>Engage domain experts and community members in GPT model design and evaluation via surveys and focus groups; and incorporate their contextual input for data preparation, prompt engineering, and model training.</td>
</tr>
</tbody>
</table>
Domestic traffic is defined as the interstate and intrastate trade of enslaved men, women, women and children. Similar to runaway ads and committal notices, domestic traffic ads were a means of communicating to the general public the subscriber’s desire to buy or sell a slave(s). Ads could be placed by private slave dealers and agents, gentry in need of domestic help, yeomen in need of extra field hands, or a public sale of an estate by the orphan’s court.” MSA
Use DTA with LLM to create a Contextualized AI Chatbot (RQ1)

PILOT FINDINGS:
DTA content was not AI-ready and needed significant pre-processing to lend itself to generative AI. GPT models need to understand the context of the DTA collection and be further trained. GPT models have parameters that allow for some control of trustworthiness. GPT models seem to be able to duplicate some of the level of analysis seen in non-AI exploratory case studies.

Perform Data Analysis on DTA dataset & Comparative Analysis (RQ2)

LangChain CSV Agent

VS.

RQ3 - How can we incorporate socio-technical considerations to promote trustworthiness and mitigate potential bias arising from the use of GPT models with library and archival collections?

- Involve user communities throughout design and development pipeline
- Capture user communities’ expectations, concepts, relationships concerning the archival collection and use of gen AI
- Incorporate feedback through regular workshops (human-human)
- Implement feedback receiving mechanisms from the chatbot (machine-human)
- Implement solutions and guardrails in place to address ethics, privacy, and cultural sensitivity
Ontology/Metadata Exploration in Our Research Design (1)

**Phase 1: GPT Selection & Focus Group Input for Model Design**

- Capture user expectations of usability and benefits; establish baselines for risks or vulnerabilities (e.g. racial or gender bias, historical or cultural inaccuracies, privacy, etc.)

**Ontology Development/Specification**

- User/User Role
- User Expectations
- Use Case
- Contextual Relevance
- Anticipated Benefits
- Risks/Vulnerabilities
- Usability
- Accessibility
- Interpretability
Ontology/Metadata Exploration in Our Research Design (2)

Phase 2: – Data Preparation & Model Training

• Explore operationalization of focus group expectations and concerns as metadata, features, or functionality to be used in the initial data preparation and/or model training. ➔ METADATA SPECIFICATION & USE IN PRE-PROCESSING FOR AI-READINESS

• Metadata added during pilot (record)
  (1) Terms of Service,
  (2) Trade Reason
  (3) Features
  (4) Terms of Sale
  (5) Owner
  (6) Optical Character Recognition (OCR) text of each DTA ad image.

• Examples of potential metadata developed through stakeholder participation (context/use/trustworthiness focus)

  • Metadata for specific use cases (e.g. historical research, K-12 education, higher education, look-up services in libraries or archives, interactive cultural heritage displays in museums, personal and professional genealogy, etc)
  • Historical contextual information
  • Labels reflecting evolution of knowledge Culturally aware labels
  • Labels reflecting user sensitivities or risk tolerances
Ontology/Metadata Exploration in Our Research Design (3)

Phase 3: Focus Group Working Meeting

- Exploring and evaluating the usability, contextual accuracy, cultural bias, and use case relevance of the chatbot ➔ REFINING ONTOLOGY & METADATA SPECIFICATION

Phase 4: GPT Performance Analysis & Re-tuning

- Targeted adaptation and retuning with respect to data annotation, metadata, model features, and model functionality, based on stakeholder feedback ➔ REVISED METADATA FOR DATA PROCESSING & MODEL RE-TUNING

Phase 5: Focus Group Working Meeting II

- Real-time evaluation of usability, contextual accuracy, bias mitigation, and use case relevance; identify technical and governance needs; areas for future research. ➔ EVALUATION: PARTICIPATORY APPROACH/SPECIFICATION/TECHNICAL
• **Status:** Pending funding decision

• **Opportunities**
  - “Strengthen the ability of libraries and archives to serve the public and research communities by unlocking new dimensions of access, discovery, and understanding within their collections.”
  - Support FAIR (Findable, Accessible, Interoperable, and Reusable) principles and reproducible computational research (RCR)
  - Prototype participatory methods that may have more general application (the broader AI community can learn from library/archive practices)
  - Contribute to generalizing use of ontologies that incorporate stakeholder participation

• **Challenges**
  - Adopting/adapting methods for human-human along with human-tech feedback
  - Operationalizing stakeholder feedback in data processing and/or model functionality – limitations? conceptual integrity?
  - Potential complexity and contradiction in stakeholder input
  - Misuse of GenAI to obscure the cultural record
We welcome your comments, discussion, and questions.

Contact Information:
Lori Perine  Lperine@umd.edu
Richard Marciano  marciano@umd.edu