DCMI 2024

Using AI to Facilitate Discoverability and Curation of ASU Library Repository Collections

DCMI Conference: Tuesday, October 22, 2024

Toronto, Canada

Arizona State University

Purpose: To investigate the suitability of ChatGPT to facilitate discoverability and curation of digital collections within the ASU Library repository ecosystem. Our project poses the question of whether ChatGPT can positively impact our ability to generate accurate metadata that aligns with relevant best practices.

What's in this presentation

- Introduction
- Project Description and Goals
- Exploration and Testing
- Initial Findings
- Future Work

Today's presenters

Timothy Provenzano

Digital Initiatives Librarian

Rachel Fernandez

Research Data Reproducibility Librarian

Chad Deets

Metadata Librarian

Deirdre Kirmis

System Architect

Project Description and Goals

ASU AI Innovation Challenge

Arizona State University

This project grew out of ASU Library's participation in the Arizona State University's AI Innovation Challenge, a university-wide initiative to grant free ChatGPT Enterprise licenses to ASU faculty, staff, and students with ideas to positively impact the future of education.

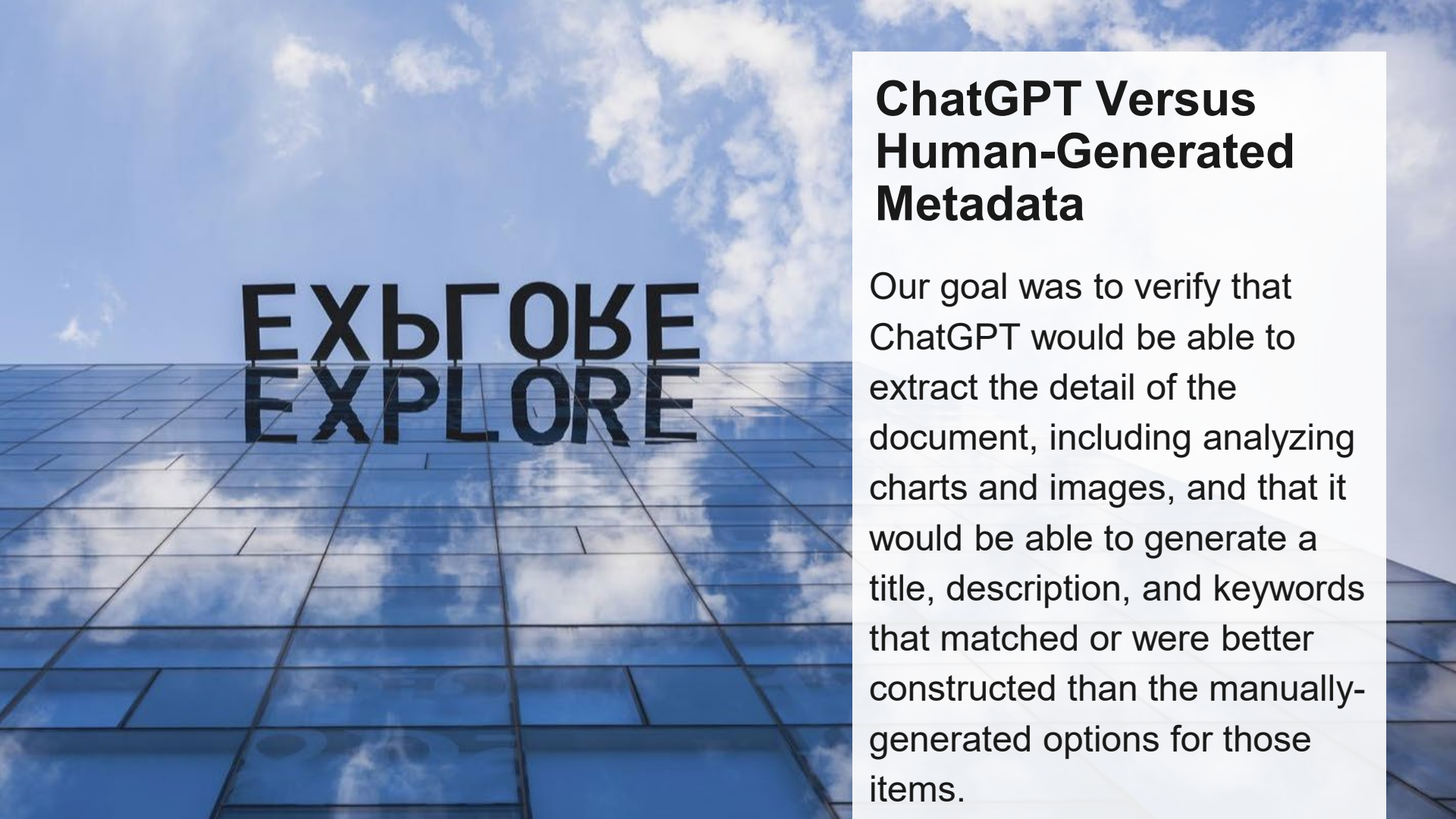
Exploration and Testing

Arizona State and Local Government Documents (SALAD) Collection

Published by:

- State of Arizona
- County and municipal governments
- Affiliated councils and agencies

http://hdl.handle.net/2286/R.C.16



ChatGPT Versus Human-Generated Metadata

Our goal was to verify that ChatGPT would be able to extract the detail of the document, including analyzing charts and images, and that it would be able to generate a title, description, and keywords that matched or were better constructed than the manually-generated options for those items.

Initial Findings

How did AI Compare ?

Titles

Often generated correct or nearly-correct titles, depending on whether title and filename were omitted in prompt.

Description

Often more concise and descriptive, unless it could not read OCR or image content. Detail depended on whether the original description or filename was omitted.

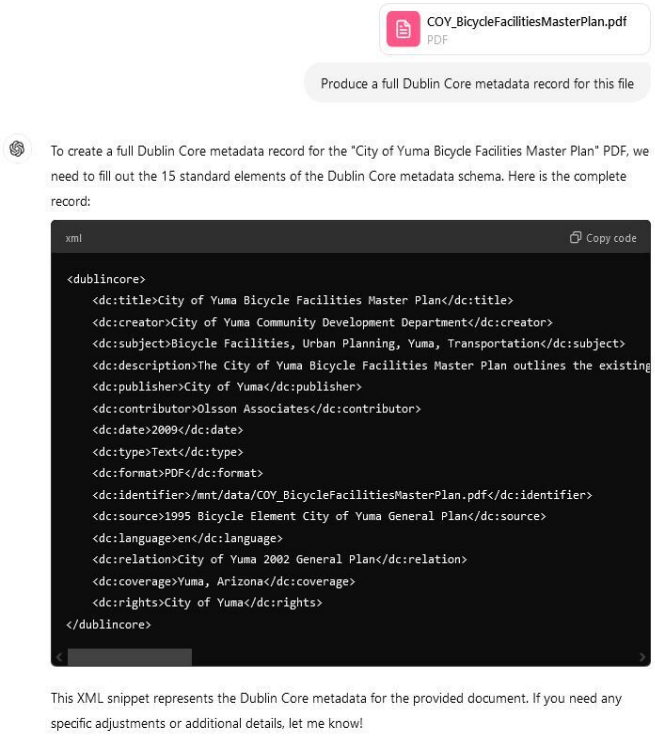
Keywords

AI suggestions similar to user-supplied terms (some missing, others added). Helped by additional prompts to pull from LCSH.

Figures

Did not extract from figures, graphs, or maps, unless captioned. Extracted text from some images, depending on how they were created and imported. Could generally read and analyze tables.

Initial Conclusions/Considerations



Curator oversight needed

Discrepancy disposition will require curator oversight level, which could become unmanageable if too numerous. Could create additional backlogs if not automated.

Source Collections

It is not anticipated that staff who will be doing the initial manual checking will need extensive curatorial expertise in the source collections. However, plausible yet meaningfully incorrect information may be harder to detect.

Enable use of paraprofessional/ student staff

Our pilot was in a restricted environment where the data could not be used to train models; however we will need to monitor what terms and conditions are placed on ChatGPT and any other tool going forward to ensure this remains the case.

Future Work

Considerations for Scaling

We hope to automate the process of metadata generation and incorporate AI within our workflows at the ASU Library.

Analyzing Multiple Files

- Extract text & summarize
- API calls cycle through files
- Tool to integrate with API & GPT

Custom GPT Training

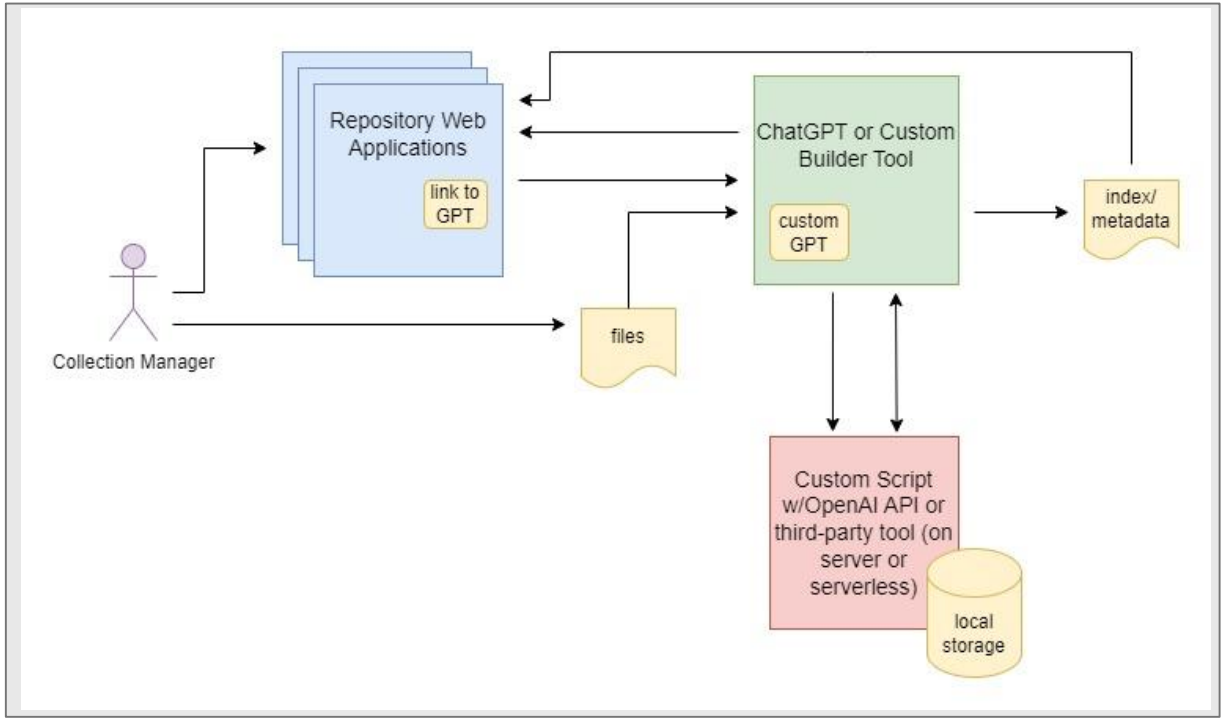
- Analyze images, graphs, maps
- Specific metadata outcomes per use case

Custom AI Model

- Custom Python program
- API calls cycle through files
- Results returned as CSV

Development and scaling with API, GPT and other tools

- Trained Custom GPT
- Local storage for multiple files
- Custom model for advanced workflow
- Utilization of API or third-party tool
- Index of file results



Thank you!

Questions?

Timothy Provenzano
timothy.provenzano@asu.edu

Rachel Fernandez
rachel.fernandez.1@asu.edu

Chad Deets
chad.deets@asu.edu

Deirdre Kirmis
dckirmis@asu.edu