

Vibe coding in the library:  
probing the potential of generative AI  
for creation of collection management  
and development tools

# Context



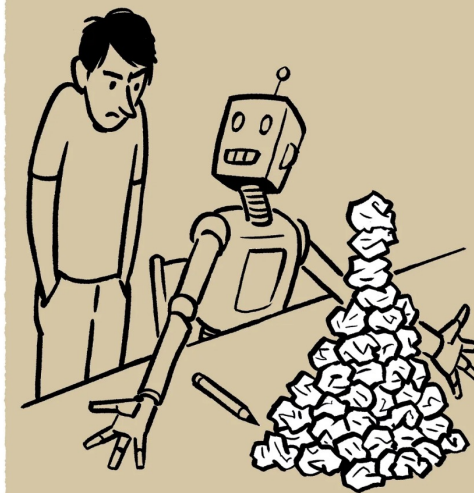
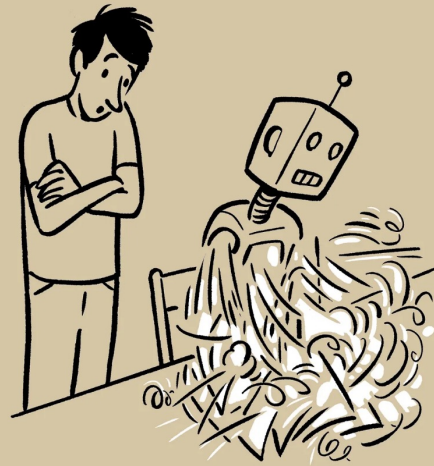
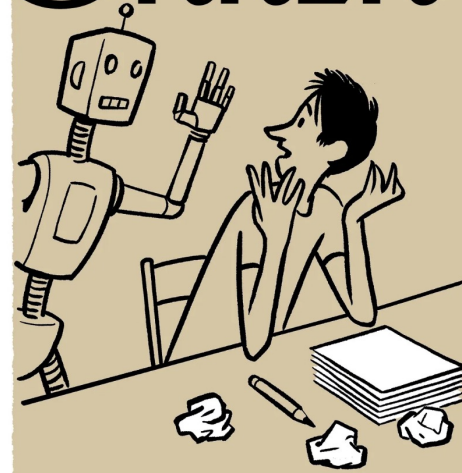
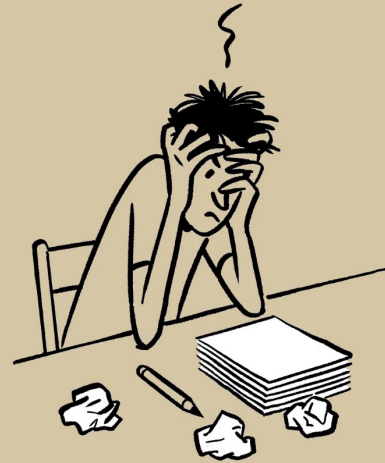
PRICE \$8.99

THE

NOV. 20, 2023

# NEW YORKER

CREATE YOUR OWN COVER ———>



CN23

## Shelfreader app

- Area: collection management
- Purpose: stocktaking, identifying missing or out-of-place books
- Status: in testing

## BibSelector tool

- Area: collection development
- Purpose: trawling widely for purchase candidates, filtering out irrelevant titles
- Status: proof of concept

# Shelfreader

## Need:

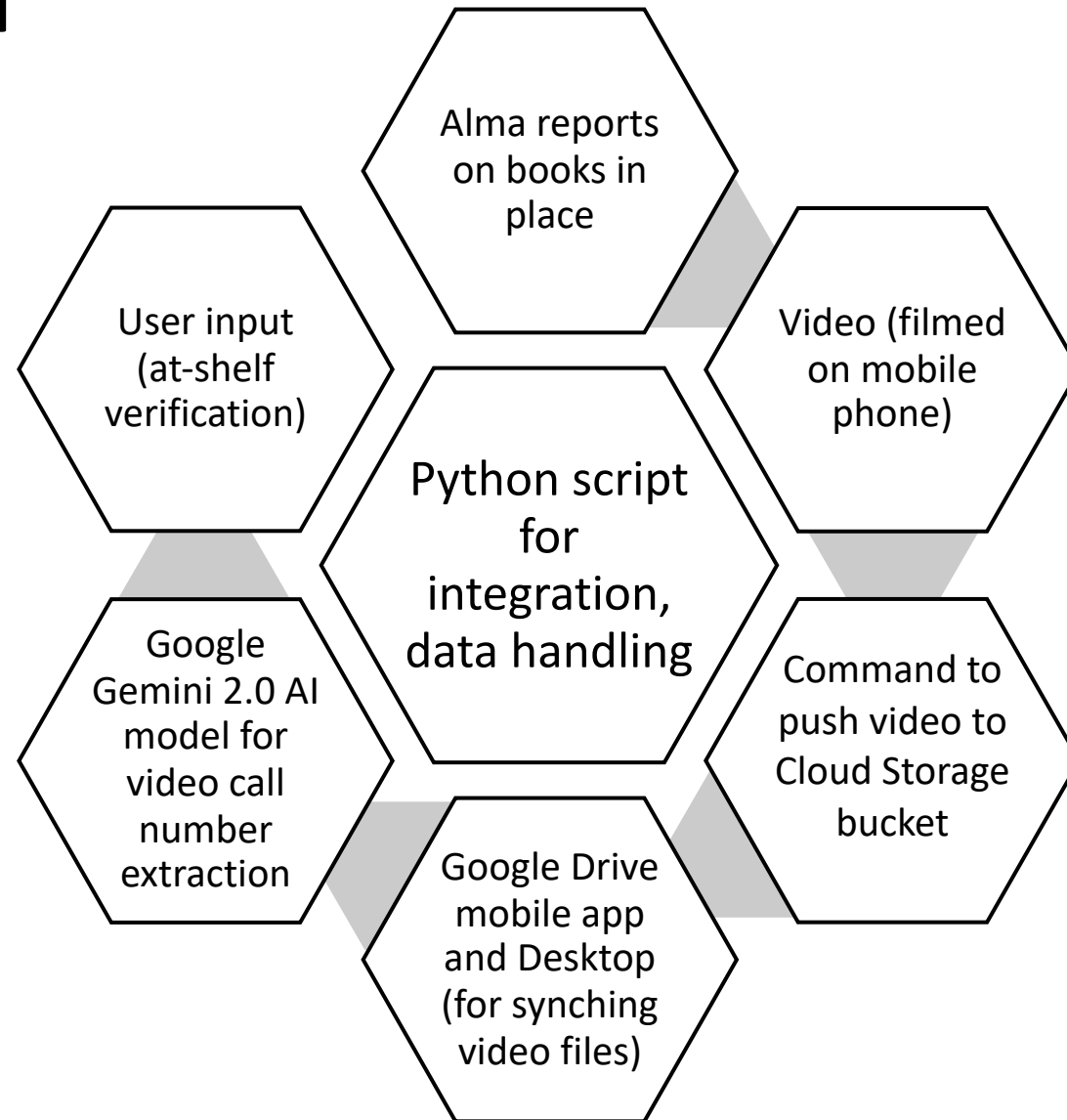
- Identifying and resolving mis-shelved books
- Identifying missing or lost items

## Current strategies:

- Manual shelfreading (infrequent)
  - No RFID system

# Shelfreader

## Components



# Sample video input

# Gemini prompt

Please forget any spine labels you have recognised before. In this video there are books in the video with visible labels on the spine and some books with no visible call number labels, usually because the books are too slim for the call number label to be placed on the spine. Although the video does not capture the entire spine of the books, you can usually tell that one book spine ends and another begins because there is a vertical line of shade. Where the books do have call number labels, you will identify the call number labels near the bottom of the spine of the books, which are rectangular white adhesive labels with rounded corners. Sometimes the corners are not visible because they wrap around the spine because the spine is narrow. Where the spines are narrow you should take extra care to make sure that you are accurately recognising where a spine label begins and ends. Even if there is not an area of contrast around a label because the books has a narrow spine you should see a vertical line of shadow which demarcates the limit between one spine label and another. Each label and its text should be recognised carefully and recorded individually. Within each label, extract the text from top to bottom across all lines, usually there are four or five lines of text. Please be careful to keep text from the same label together. The call numbers are Library of Congress Classification call numbers. The first line almost always contains a letter. The second line almost always contains a numerical value. The third line generally contains a combination of punctuation, letters and numbers, e.g. ".R26" . The fourth line can contain a date, e.g. 1900, or another alphanumerical value, e.g. R66. If there is an alphanumerical value in the fourth line, the fifth line is usually a date, e.g. 1900.



# Shelfreader

# Shelfreader

## Evaluation

- Time savings?
- Accuracy
- Different affordances:
  - Snapshot of the shelves as they are at a given moment
  - Provides granular mapping of the books on each shelf, in each section
    - useful for signage creation
- Deploy now or wait for incremental LLM improvements?

# BibSelector tool for collection development

Need:

- Streamline selection for unusual collecting profile
  - geographically and chronologically limited
  - spanning all disciplines, LC classes

Current strategies:

- reliance on periodic manual keyword searches, reports
- purchase requests
- approval plans

# BibSelector tool for collection development

## Collecting summary

### Italy 1200-1650

As its primary collecting focus, the Berenson Library comprehensively acquires works of current scholarship in all subjects related to the region conventionally called “Italy” in the timeframe roughly between 1200 and 1650, in all relevant languages and formats. Older but still important works are collected when they are unavailable in electronic format. ... In our period, the Italian peninsula contained many distinct city-states or principalities, some of which were at times under the rule of non-Italian powers such as Spain or France. But “Italy” in this period also included areas outside the peninsula that were then ruled by or under the influence of Italian states, such as the maritime empire of Venice, which extended down the Balkan peninsula and far into the eastern Mediterranean. Studies on the historiography of the Italian Renaissance also fall into the library’s primary collecting scope. Every LC class and most subclasses are represented, and the large majority of our acquisitions are concentrated in this area. The Berenson Library has several secondary areas within its collecting scope, almost all of them associated closely with the field of Italian Renaissance studies. They include the following:

### Renaissance period outside of Italy (1200-1650)

In this area the library collects selectively in a broad geographical range that concentrates on Europe, encompasses the entire Mediterranean basin, and extends further afield to the New World and to South and East Asia. ...

**Medieval studies [...]**

**Music history [...]**

**Classical studies [...]**

Original pool of acquisition candidates 36435

Preprocessing +  
Keyword Filters

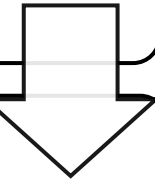
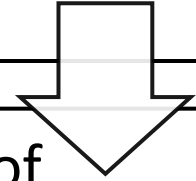
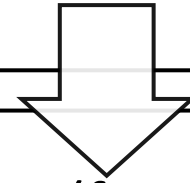
Filtered pool of acquisitions candidates (first  
pass) 16315

Semantic Similarity Scoring

Filtered pool of  
candidates (second pass)  
9547

LLM evaluation against collection development  
policy

Pool for human  
verification 3184  
(1337 'yes' only)



## Preprocessing

- Exclude if no subject headings present
- Data is insufficient for analysis

## Filtering by keywords and dates

- Exclude by subject heading keywords
- Exclude by chronological subdivisions
- Reintegrate any candidates that contain key phrases

✗ "Juvenile literature"  
✗ "Novels"

?

Date range patterns:  
Excluded if no overlap  
1200-1650

✓ "Renaissance"  
✓ "Early works to 1800"  
✓ "Historiography"  
✓ "16th century"

## Embedding s

- Sentences (titles, subject headings) transformed into embeddings
  - books already purchased (historical acquisitions, 2015-2023)
  - candidates for future purchase (books acquired by Harvard libraries, 2024- )

## Scoring

- Semantic similarity scoring
- Similar sentences appear close together in vector space

## Best match

- Highest similarity match between historical acquisitions and acquisitions candidates calculated
- Candidates with similarity matches under 0.65 are excluded

## Evaluation by LLM

- API requests to Claude 3.7 Sonnet containing:
  - Bibliographical information
  - Collection development policy
  - Prompt

## Filter by decision

- Exclude "No" category
- Evaluate "Yes" and "Maybe" categories manually

Prompt: "You are an acquisitions librarian evaluating books for purchase. Applying the criteria in the collection development policy, provide a purchase recommendation (yes/maybe/no) for each title and explain your decision."

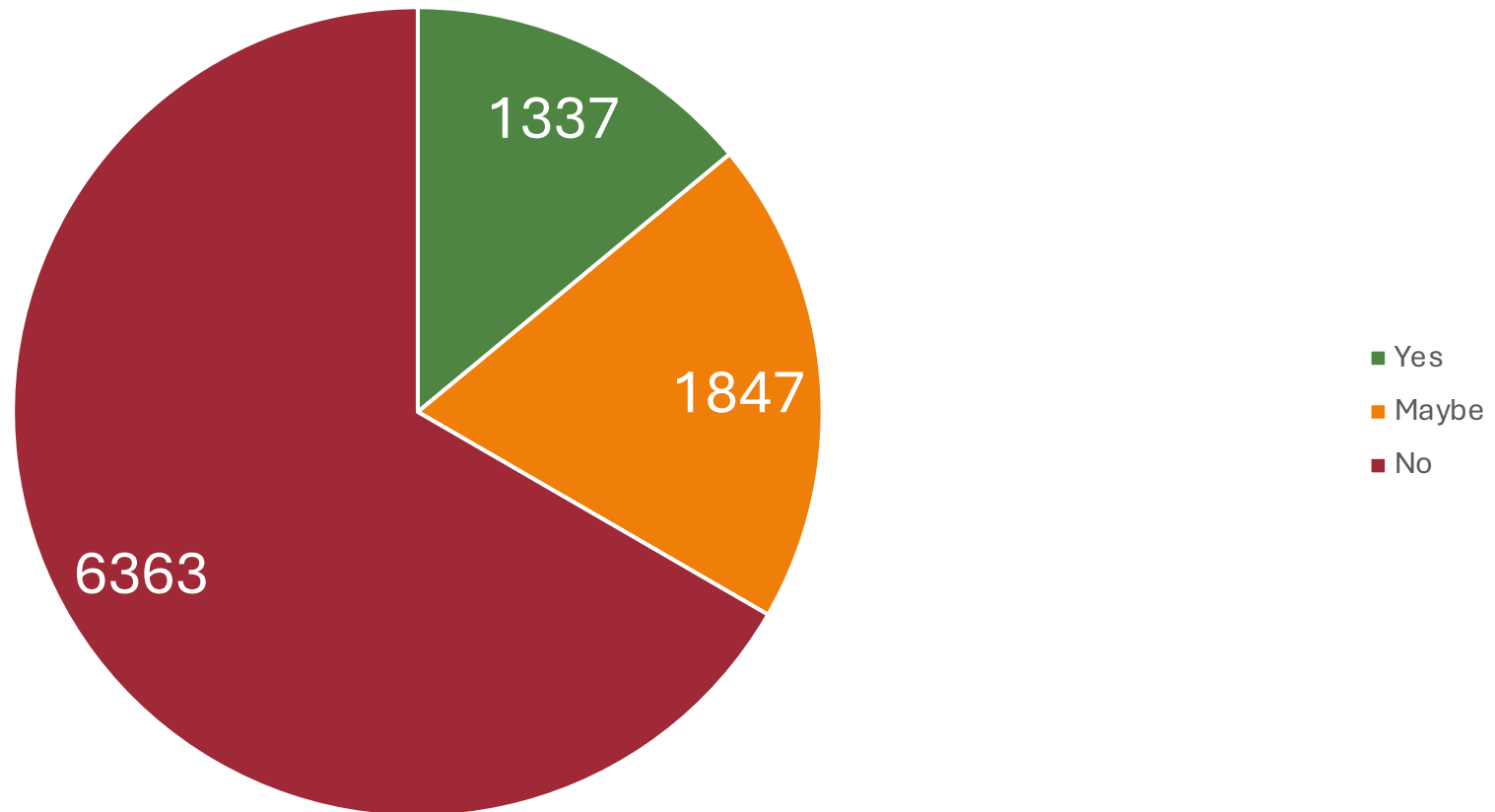


# Validation

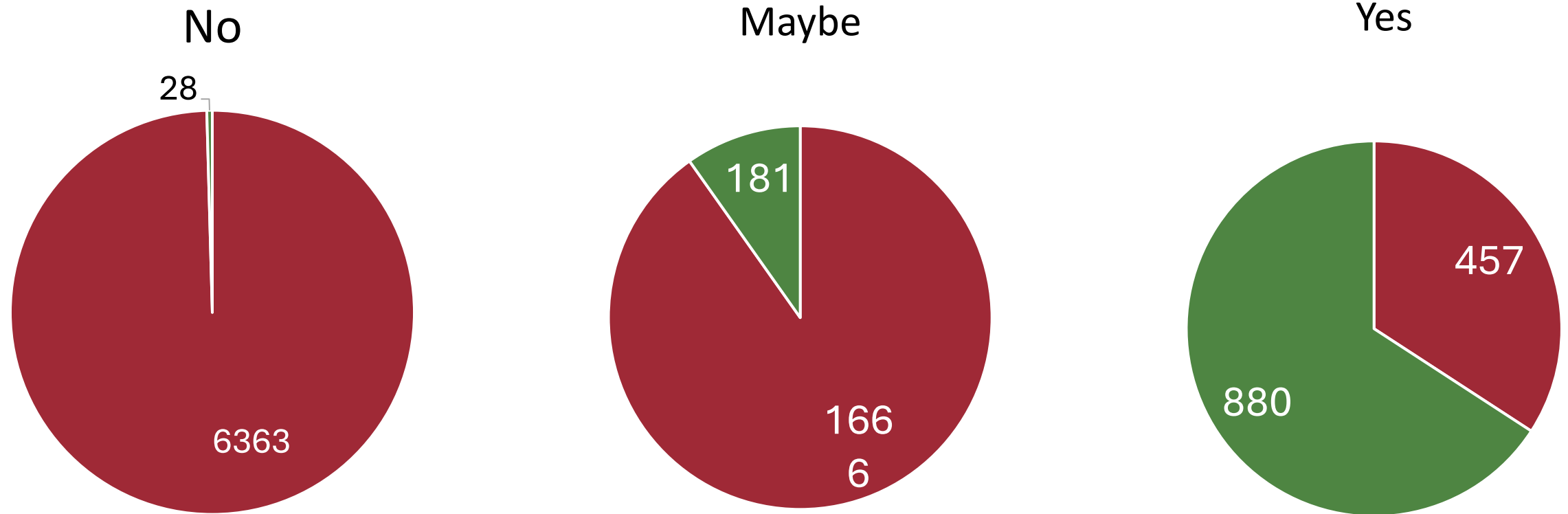
How accurately did the tool predict which books added by Harvard Libraries after 2023 would be purchased by Biblioteca Berenson?

- Human verification – do recommendations look right?
- Comparison with real-life acquisitions

# Purchase recommendations: Yes/No/Maybe



# Recommendations vs. reality



Not acquired by Biblioteca Berenson

Acquired by Biblioteca Berenson

# BibSelector

## Evaluation

- Exclusion of records seems to reflect collecting practice well
- Successfully retrieves some relevant but unacquired books for evaluation
- Incompleteness of bibliographical data (especially abstracts, tables of contents) is a significant constraint on what tool can achieve
- What mechanisms would be required to make sure tool accommodates evolution in research interests and collection practices?

# BibSelector

Next steps:

- Continuing evaluation of results
- Evaluation of benefits of data enrichment by Perplexity AI
- Further integration and automation with other systems and data sources to improve efficiency further
- Testing with other sets of bibliographical records

# Conclusions

Extra slides

# Tools for scripting

## Models:

- ChatGPT (OpenAI)
- Claude (Anthropic)
- Gemini (Google, available in Colab)
- Copilot (Microsoft)
- Ollama models (Meta, can be run locally)

## Scripting environments:

- Cursor
- VS Code with Cline
- Windsurf



# Financial costs

## Claude 3.7

- Input: \$3.00 per million tokens
- Output: \$15.00 per million tokens
- 50% discount for batches

## Perplexity (due to rise in April)

- \$20.00 per month **plus:**
- Input: \$3.00 per million tokens
- Output: \$15.00 per million tokens
- \$5.00 per 1000 searches

## Gemini 2.0

- Input: \$0.10 (text / image / video)
- Output: \$0.40 per million tokens

## Cursor

- \$20.00 per month

# Environmental costs

- Energy and water use poorly understood, little information released
- Some attempts have been made to simulate using open source models
- Complex to estimate (cost of running models vs cost of training)
- Good practice:
  - Run offline where possible
  - Triage so that the most computing-intensive processes are only run where necessary
  - Prepare well and run once

# Resources for self-guided learning

- Colleagues
- Youtube – Matthew Berman
- Online tutorials (Medium and elsewhere)
- Staff skills-share programs
- Mailing lists:
  - Harvard Baker Library's generative AI newsletter
  - One Useful Thing
- Conversations with Claude 3.7, LLM of your choice
- Get stuck in