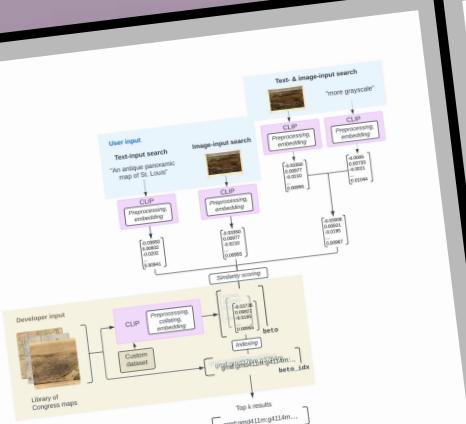


CHR2024

Best Long Paper Award

awarded to

Jamie Mahowald & Benjamin Lee



Integrating Visual and Textual Inputs for Searching Large-Scale Map Collections with CLIP

Jamie Mahowald¹, Benjamin Charles Germain Lee^{2*}
¹The Archer Center & Department of Mathematics, The University of Texas at Austin, USA
²Information School, University of Washington, USA

Abstract
Despite the prevalence and historical importance of maps in digital collections, current methods of navigating and exploring map collections are largely restricted to catalog records and structured metadata. In this paper, we explore the potential for interactively searching large-scale map collections using natural language inputs ("maps with sea monsters") and visual inputs (i.e., reverse image search), and multimodal inputs (an example map + "more grayscale"). As a case study, we use 1462,842 images of maps publicly accessible via the Library of Congress's API. To accomplish this, we use the multimodal Contrastive Language-Image Pre-training (CLIP) machine learning model to generate embeddings for these maps, and we develop code to implement exploratory search capabilities with these different strategies. We present results for example searches created in consultation with staff in the Library of Congress's Geography and Map Division, and describe the strengths, weaknesses, and possibilities for these search queries. Moreover, we introduce a fine-tuning architecture of 10,504 map-caption pairs, along with an architecture for fine-tuning CLIP models on this dataset. To facilitate re-use, we provide all of our code in documented, interactive Jupyter notebooks and place all code into the public domain. Lastly, we discuss the opportunities and challenges for applying these approaches across both digitized and born-digital collections held by galleries, libraries, archives, and museums.

Keywords
maps, Library of Congress, computing cultural heritage, multimodal machine learning, exploratory search

5. Results & Discussion

In this section, we introduce example search results for all three strategies described in Section 4 and reflect on the strengths and limitations of our implementation. We also discuss the utilization of the LC Labs AI Planning Framework throughout our research process.

