Searching Digital Heritage: Putting IR Evaluation in Context

Jaap Kamps

Faculty of Humanities, University of Amsterdam
kamps@uva.nl

Anyone offering cultural heritage content in a digital library is naturally interested in assessing its performance: how well does my system meet my searchers' information needs? Standard evaluation benchmarks in the Cranfield/TREC paradigm allow us to research the generic retrieval effectiveness of system, by abstracting away from the specific document genre, use-case and searcher stereotype. While this is of clear value, it also ignores the unique content and user community of a particular digital library. How does this content differ, and how do the tasks and searchers differ? Can we make the evaluation tailored to their unique characteristics? Moreover, can we distinguish different types of usage (e.g., the professionals versus the public) and how does this impact the evaluation?

Keywords: Cultural Heritage, Information Retrieval, Evaluation