Influence of Navigation Leads' Visualization on Digital Libraries Exploration

Róbert Móro

Supervisor: Mária Bieliková

Motivation

- Researching a new domain
  - exploratory search task
  - researcher novice scenario
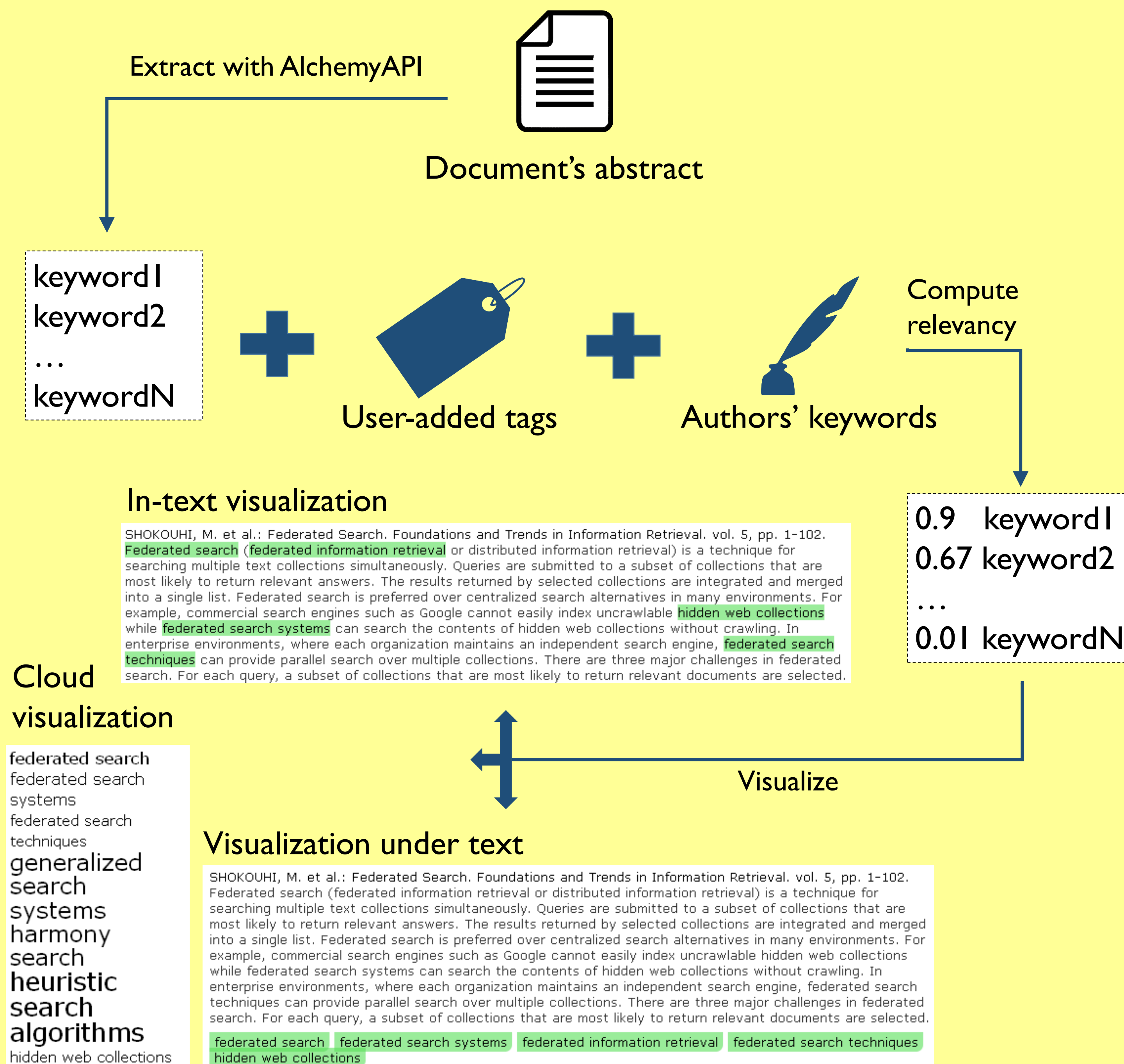- How does visualization of search interface affect the novices' task performance?

Navigation Leads

- Important words (phrases) in a summary or an abstract
  - filter relevant documents
- Relevancy based on associated metadata
  - tags, authors' keywords
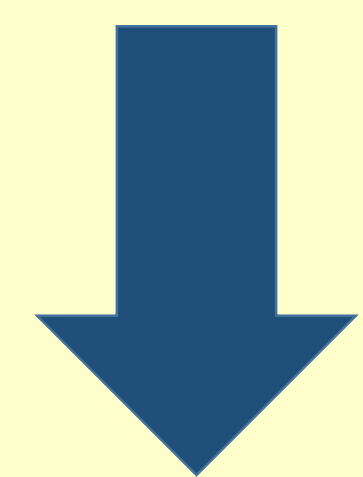 - title, abstract

Evaluation

- User study
  - 5 participants, 1+3 tasks
  - questionnaire, search logs, video, audio, eye-tracking
- Annota
  - 16 000 articles from ACM DL

Leads' Selection and Visualization

Extract with AlchemyAPI

Document's abstract

keyword1
keyword2
…
keywordN

User-added tags   +   Authors' keywords

Compute relevancy

0.9   keyword1
0.67 keyword2
…
0.01 keywordN

In-text visualization

SHOKOUHI, M. et al.: Federated Search. Foundations and Trends in Information Retrieval. vol. 5, pp. 1-102. Federated search (federated information retrieval or distributed information retrieval) is a technique for searching multiple text collections simultaneously. Queries are submitted to a subset of collections that are most likely to return relevant answers. The results returned by selected collections are integrated and merged into a single list. Federated search is preferred over centralized search alternatives in many environments. For example, commercial search engines such as Google cannot easily index uncrawlable hidden web collections while federated search systems can search the contents of hidden web collections without crawling. In enterprise environments, where each organization maintains an independent search engine, federated search techniques can provide parallel search over multiple collections. There are three major challenges in federated search. For each query, a subset of collections that are most likely to return relevant documents are selected.

Cloud visualization

federated search
federated search systems
federated search techniques
generalized search systems harmony search heuristic search algorithms
hidden web collections

Visualize

Visualization under text

SHOKOUHI, M. et al.: Federated Search. Foundations and Trends in Information Retrieval. vol. 5, pp. 1-102. Federated search (federated information retrieval or distributed information retrieval) is a technique for searching multiple text collections simultaneously. Queries are submitted to a subset of collections that are most likely to return relevant answers. The results returned by selected collections are integrated and merged into a single list. Federated search is preferred over centralized search alternatives in many environments. For example, commercial search engines such as Google cannot easily index uncrawlable hidden web collections while federated search systems can search the contents of hidden web collections without crawling. In enterprise environments, where each organization maintains an independent search engine, federated search techniques can provide parallel search over multiple collections. There are three major challenges in federated search. For each query, a subset of collections that are most likely to return relevant documents are selected.

federated search   federated search systems   federated information retrieval   federated search techniques   hidden web collections

Results

- In text
  - **+** immediate context
  - **−** too distractive when reading
- Under text
  - **+** less distractive; context of a document
  - **−** lacks immediate context
- In cloud
  - **+** differentiates leads' global relevance
  - **−** looses all context

Conclusions

- visualization under text with in-text highlighting at hover
- relevance differentiation