Meaningful Concept Displays for KOS Mapping

Xia Lin & Jae-Wook Ahn

Drexel University

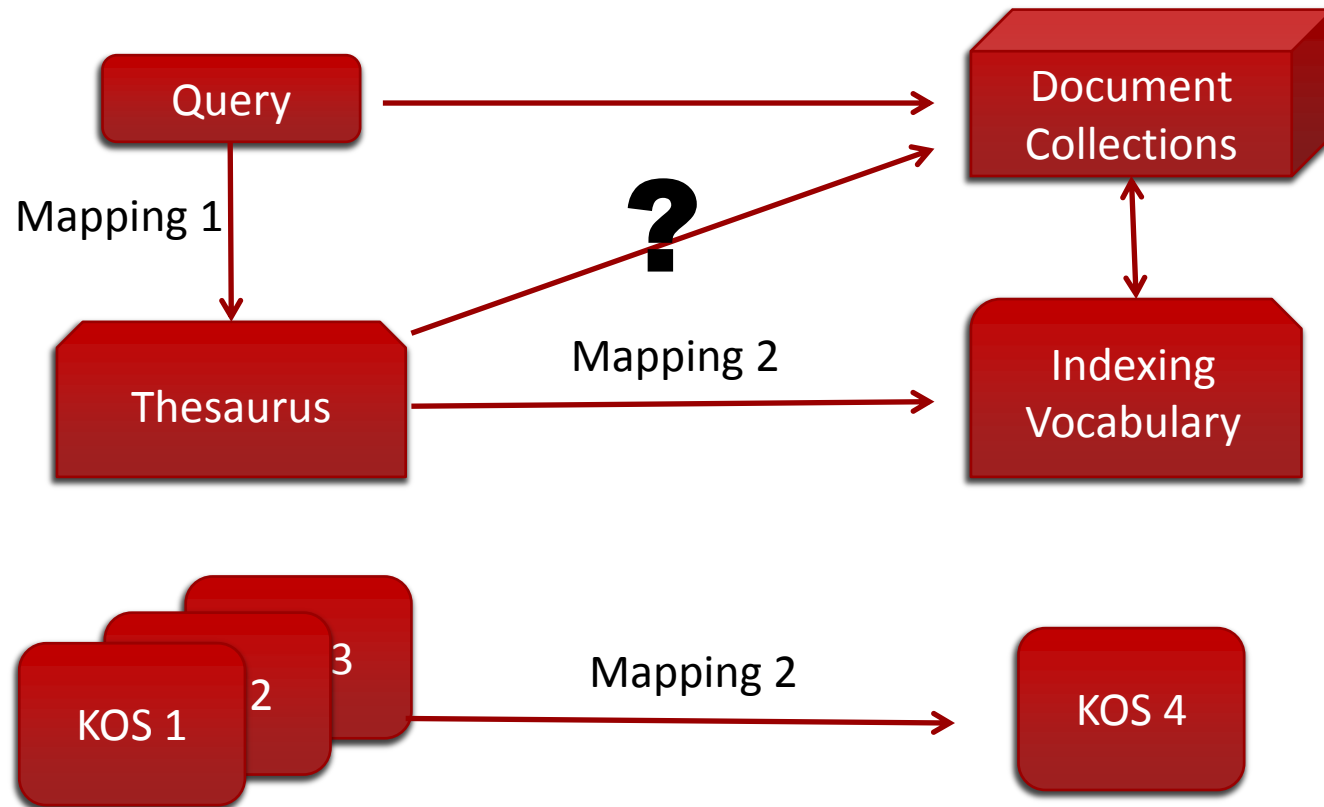Dagobert Soergel

The University at Buffalo

◆ There are increasing numbers of KOS available through Linked Data

> ➢ They are all described in RDF or OWL.

> ➢ They may be aggregated in one single endpoint and searchable through SPARQL queries

> ➢ Are they linkable at the concept level?

>> ➢ Some are controlled vocabularies, some are not

>> ➢ Some are elemental and some are compound concepts

>> ➢ Many concepts will have different scopes

◆ Linked data provide only a syntactical solution to KOS mapping. Much more work need to be done at the semantic level.

DREXEL UNIVERSITY
College of
Computing & Informatics

Old problems? New Problems?

◆ We have had many years research on

➤ Mapping free text queries to terms in controlled vocabularies

➤ Mapping from controlled vocabulary A to controlled vocabulary B ?

➤ Mapping of controlled vocabularies A and B through C

➤ Integrate multiple vocabulary resources to one (i.e., UMLS)

◆ What's new?

➤ Networked environment -- Instant connection through URI

➤ Interactive environment – Users make selections during the mapping process

➤ Integrated environment – vocabularies are linked to their primary resources.

Old Problems? New Problems?

◆ Testing KOS mapping as "query expansion"

New Challenges

◆ KOS mapping as "query expansion"

Query → Document Collections

Mapping 1

KOS 1, 2, 3 — Mapping 2 → KOS 4

Document Collections ↕ KOS 4

DREXEL UNIVERSITY
College of
Computing & Informatics
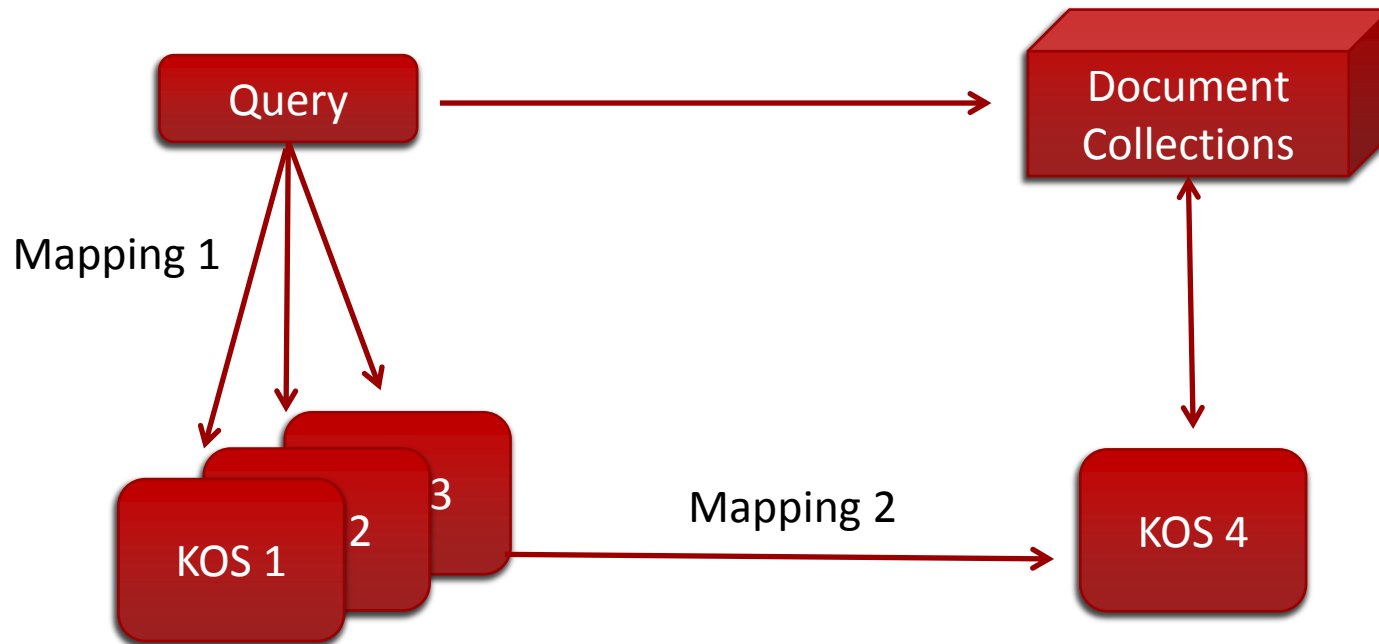
Real World Problems

◆ Desire: Use Getty's AAT to search ARTstor collections

◆ Problems: ARTstor collections are NOT indexed by AAT
  ➢ Some are indexed by their own controlled vocabularies
  ➢ Some are indexed by just a list of keywords

◆ Previous approach:
  ➢ Search the collections directly by AAT terms
  ➢ Ontology annotation – attempt to automatically assign AAT terms to ARTstor collections

◆ Our approach:
  ➢ Map the queries to AAT terms and AAT terms to indexing terms

DREXEL UNIVERSITY
College of
Computing & Informatics
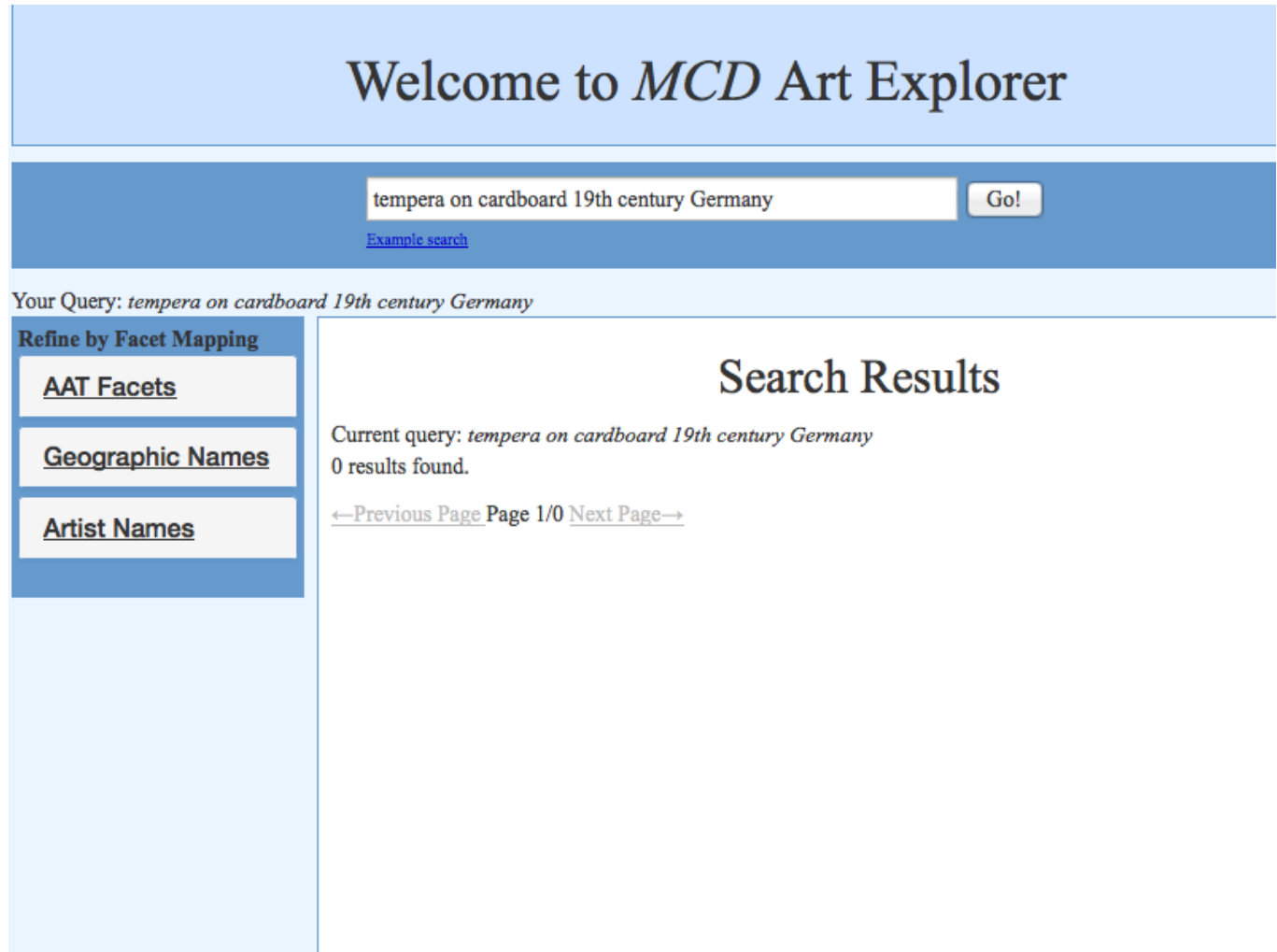
Mapping Strategies

- Decompose queries (or terms) to elementary concepts whatever possible

- Identify facets of the elementary concepts

- Map the elementary concepts to CV terms

- Map CV terms to indexing terms

- Let the users select indexing terms that match their needs

- Use the facets to narrow down search results

DEMO: Mapping and Search

Query:
"Tempera on Cardboard 19th century Germany"

Direct search get no results.

Welcome to *MCD* Art Explorer

tempera on cardboard 19th century Germany   Go!
Example search

Your Query: *tempera on cardboard 19th century Germany*

Refine by Facet Mapping

AAT Facets

Geographic Names

Artist Names
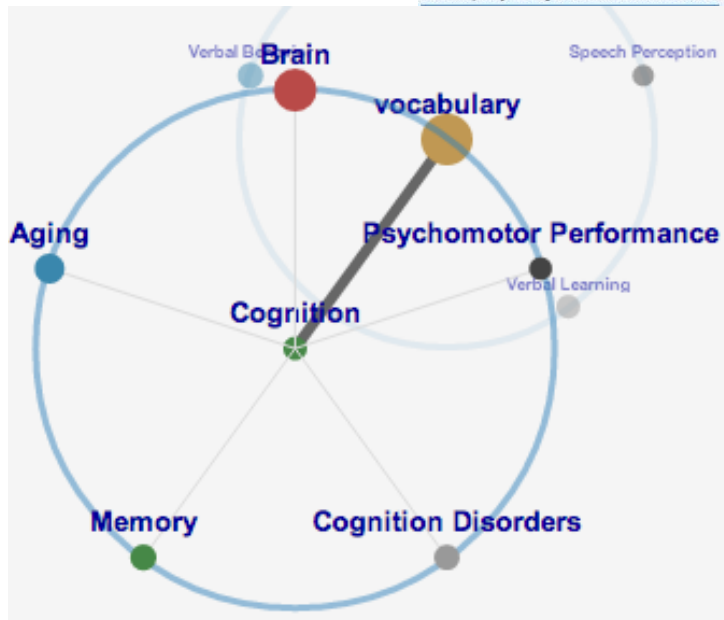
Search Results

Current query: *tempera on cardboard 19th century Germany*
0 results found.

←Previous Page **Page 1/0** Next Page→

DREXEL UNIVERSITY
College of
Computing & Informatics

Mapping demo

The user selects AAT term:
"egg tempera"
Which maps to
ARTstor terms:

- Egg tempera, gesso on wood
- Egg tempera on wood
- Egg tempera on canvas
- Egg tempera on masonite
- Egg tempera on board
- Egg tempera on panel
- …….

Mapping 1

◆ Use NLP and patterns to map the query:

◆ "tempera on cardboard 19th century Germany"
  ➢ Subject: "tempera on cardboard"
    ➢ Pattern: "painting on surface"
      ➢ Tempera → AAT terms:
        ➢ Tempera, egg tempera, gom tempera, wax tempera, ….
      ➢ Cardboard →AAT terms: …..
  ➢ Date and time: "19th century"
  ➢ Location: "Germany"

Mapping 2

◆ Map all the ARTstor indexing terms to AAT terms

➢ Which is a batch processing done during the system implementation and the results are saved in the databases

➢ During the search, the results can be showed instantly

➢ The user can click on a AAT term and see the matching ARTstor terms

➢ The search engine can display search results if any of the indexing terms are selected.

Example 2: search for "China Clay"

Initial query:
"China Clay"

-- the search is based on string matching only.

Example 2

Queries after mapping:

"Wood, Kaolin;
  Wood, Kaolin, animal hair;
  Wood, Kaolin, features"

Towards Meaningful Displays

Towards Meaningful Displays

Towards Meaningful Displays

Conclusions

◆ Meaningful displays come from meaningful mapping.

◆ KOS mapping can be done at both the global level (the whole vocabulary) and the local level (individual terms)

◆ KOS mapping needs to be done interactively
   ➢ The algorithmic mapping results are presented through MCDs.
   ➢ The user interacts with MCD to get the best of the mapping results.
   ➢ The mapping process and the searching and browsing process should be integrated.