Interactive Information Access on the Web of Data

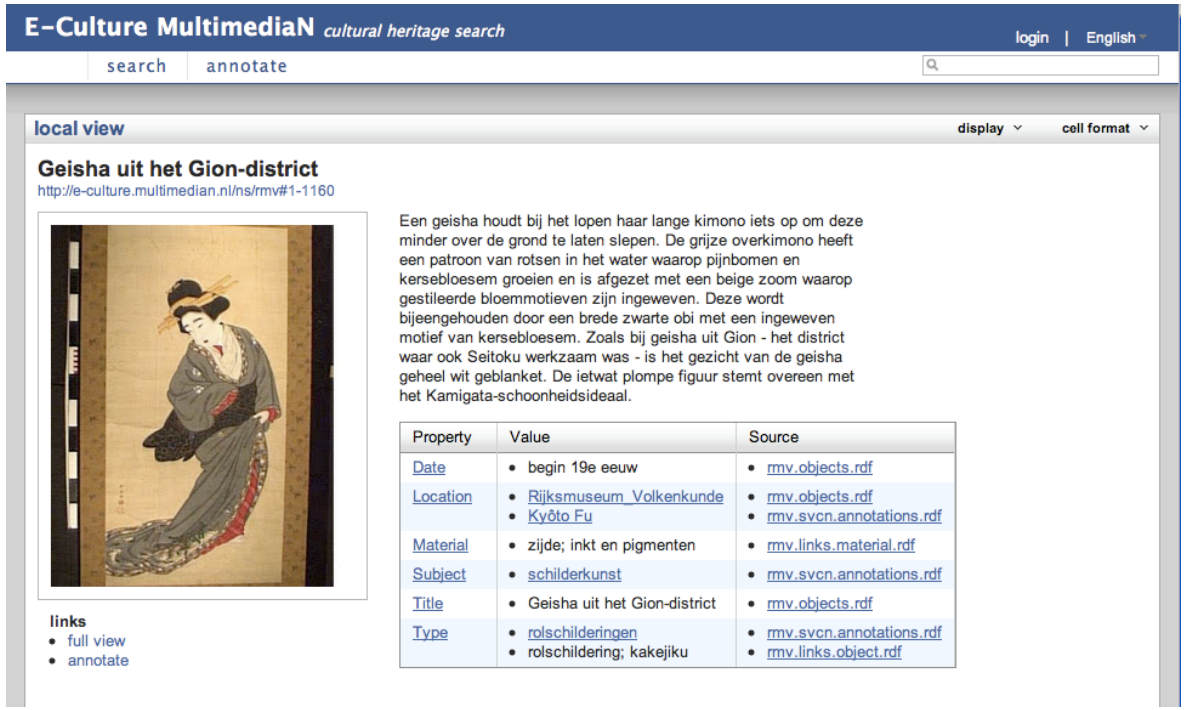
Lynda Hardman, Jacco van Ossenbruggen,
Raphaël Troncy, Alia Amin and Michiel Hildebrand

Interactive Information Access

http://www.cwi.nl/interactive_information_access

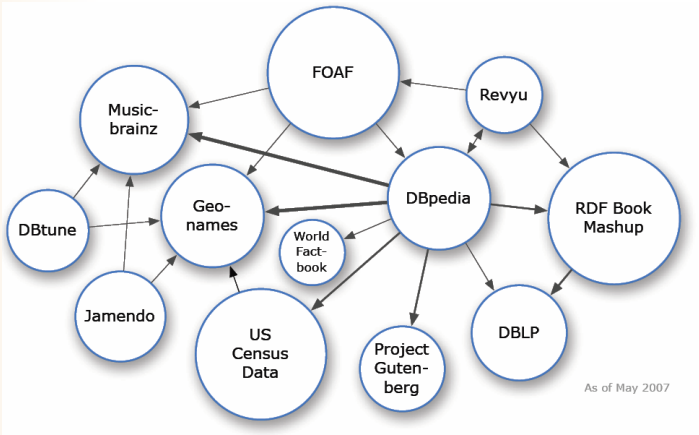
What is linked data?

- URIs, possibly identifying media fragments
- + annotations (tags)
- + links among fragments & annotations



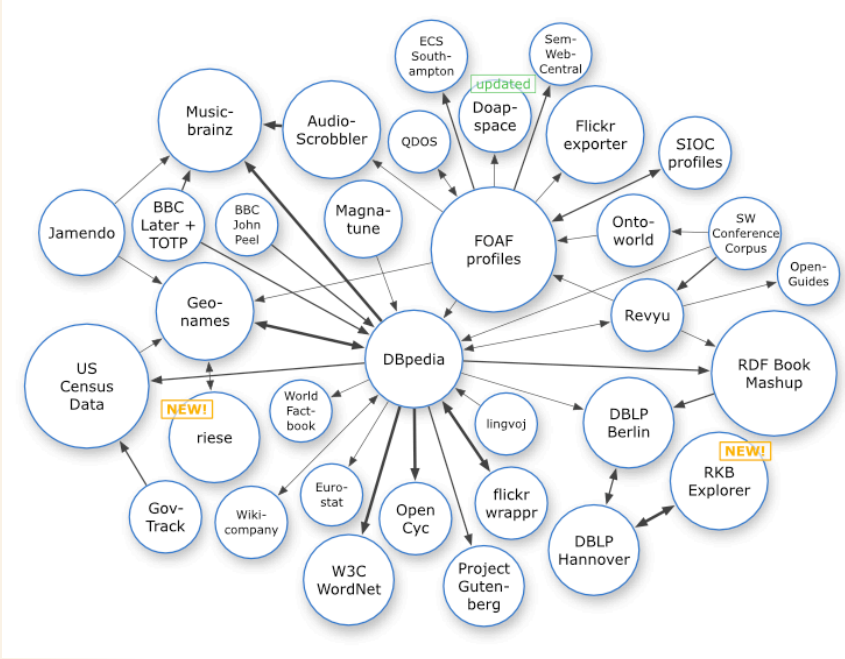
How much linked data is there?

May 2007



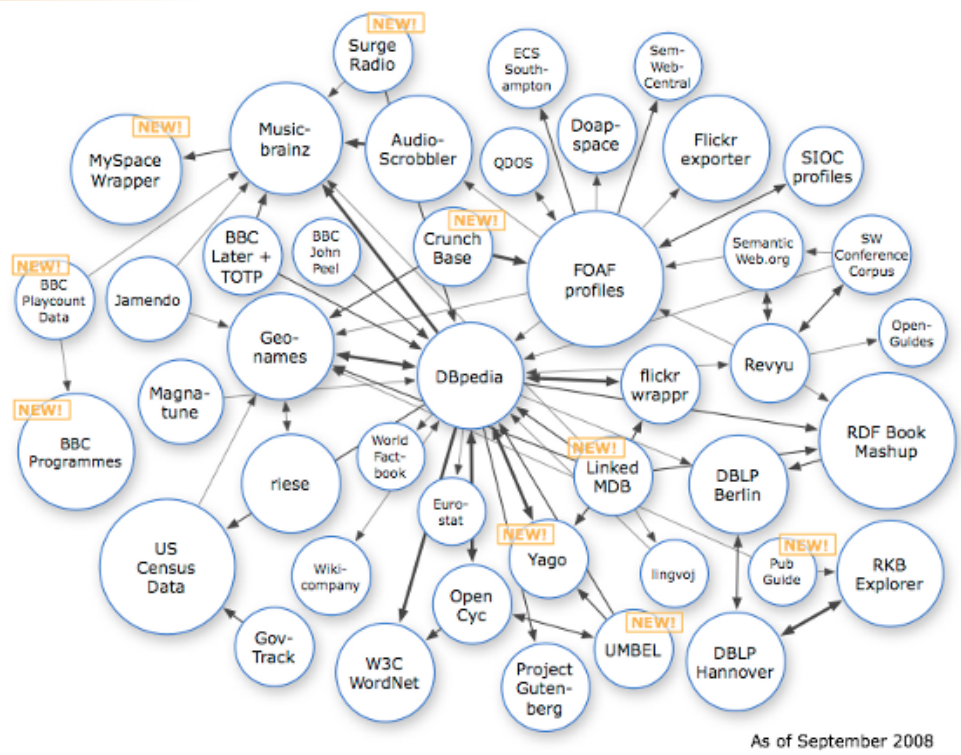
Credit: Chris Bizer

Linked data cloud March 2008



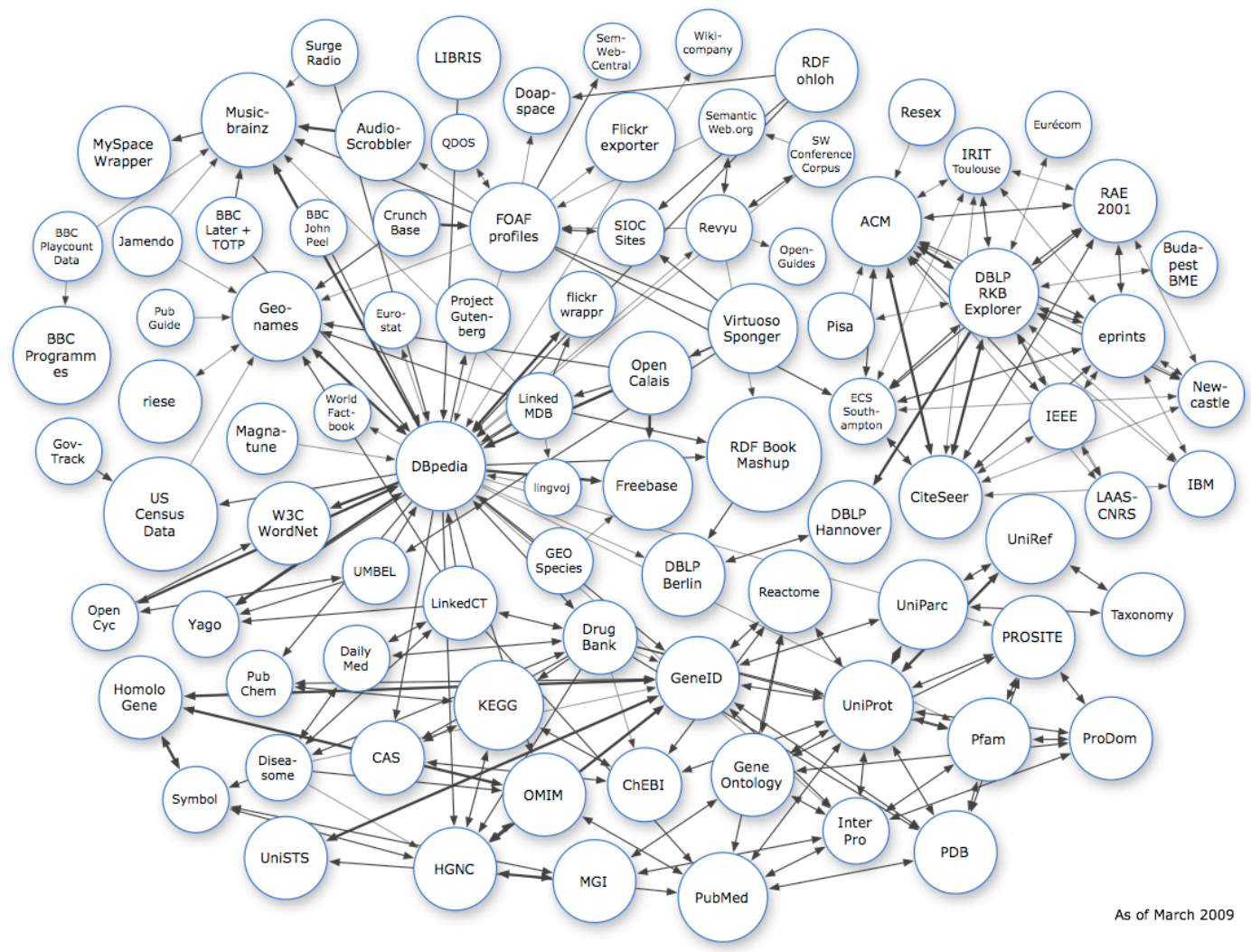
Credit: Richard Cyganiak

Linked data cloud September 2008



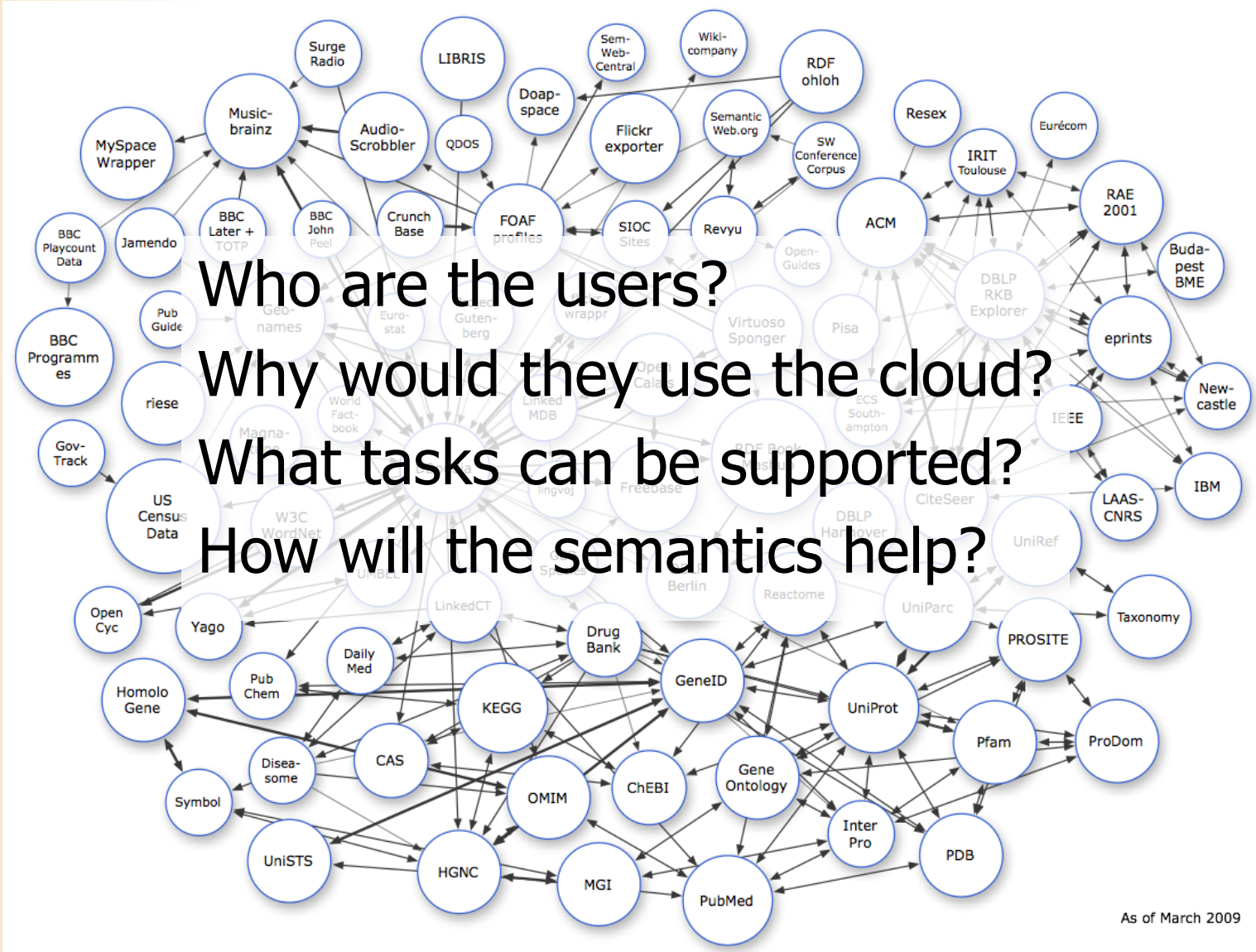
http://www4.wiwiss.fu-berlin.de/bizer/pub/lod-datasets_2008-09-18.png

Linked data cloud March 2009



> 4.5 billion RDF triples, interlinked by around 180 million RDF links

http://www4.wiwiss.fu-berlin.de/bizer/pub/lod-datasets_2009-03-05.png



Who are the users?

Why would they use the cloud?

What tasks can be supported?

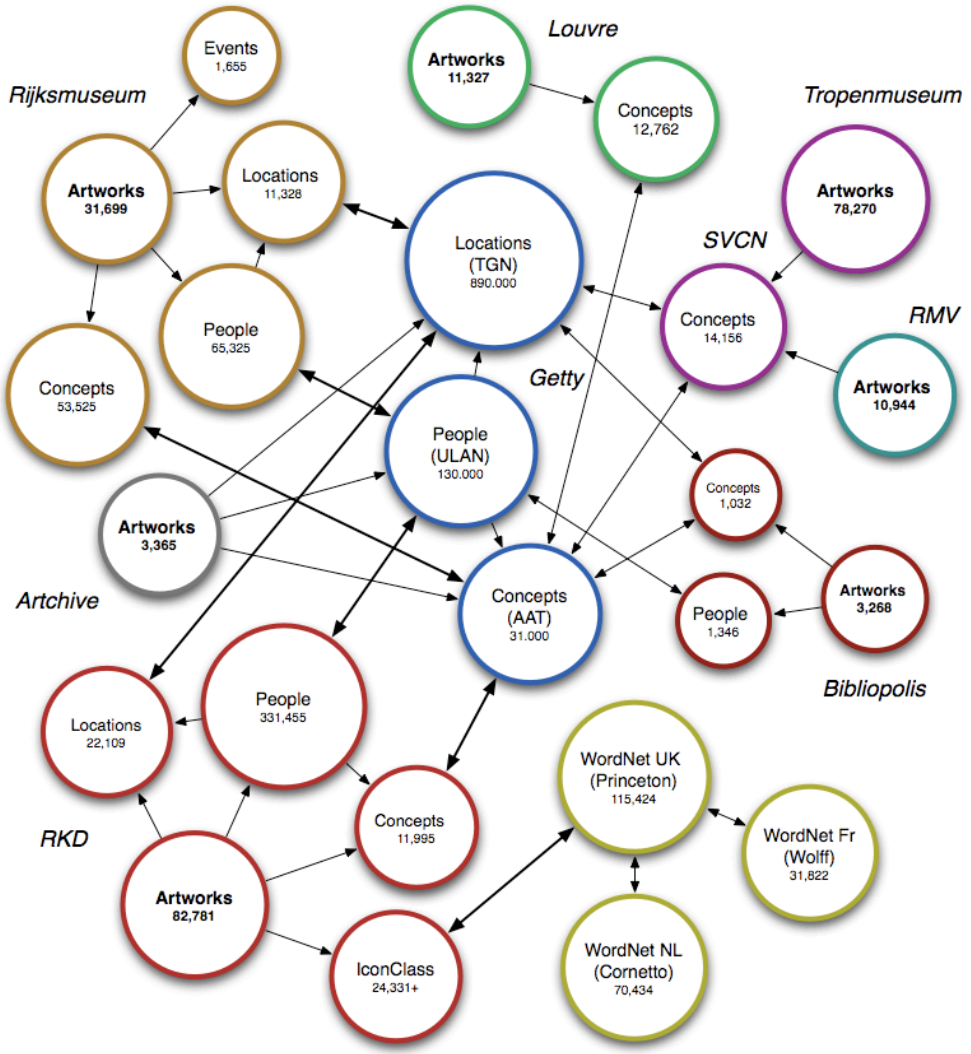
How will the semantics help?

As of March 2009

How can semantics help?

- Query construction
  - disambiguate input
  - selection of available terms
- (Semantic) search algorithm
  - graph traversal
  - query expansion
  - RDFS/OWL reasoning
- Presentation of search results
  - grouping by property
  - visualization on timeline, map

Data sets in E-Culture demo



http://e-culture.multimedian.nl/resources/datacloud/

Michiel Hildebrand
http://e-culture.multimedian.nl/

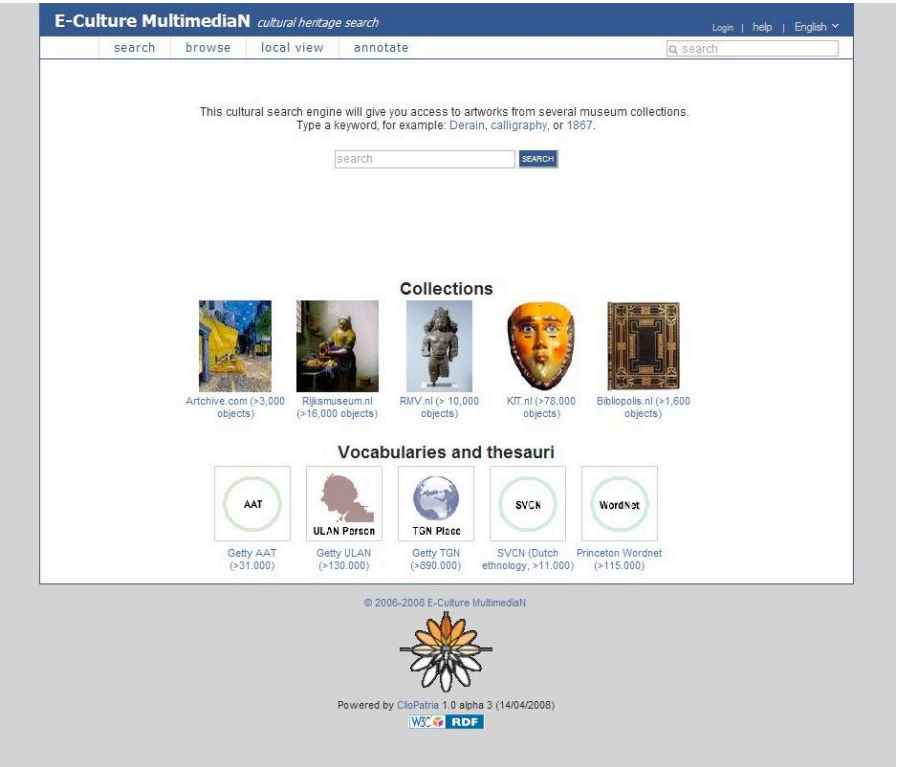
Browsing annotated collections of cultural heritage artefacts

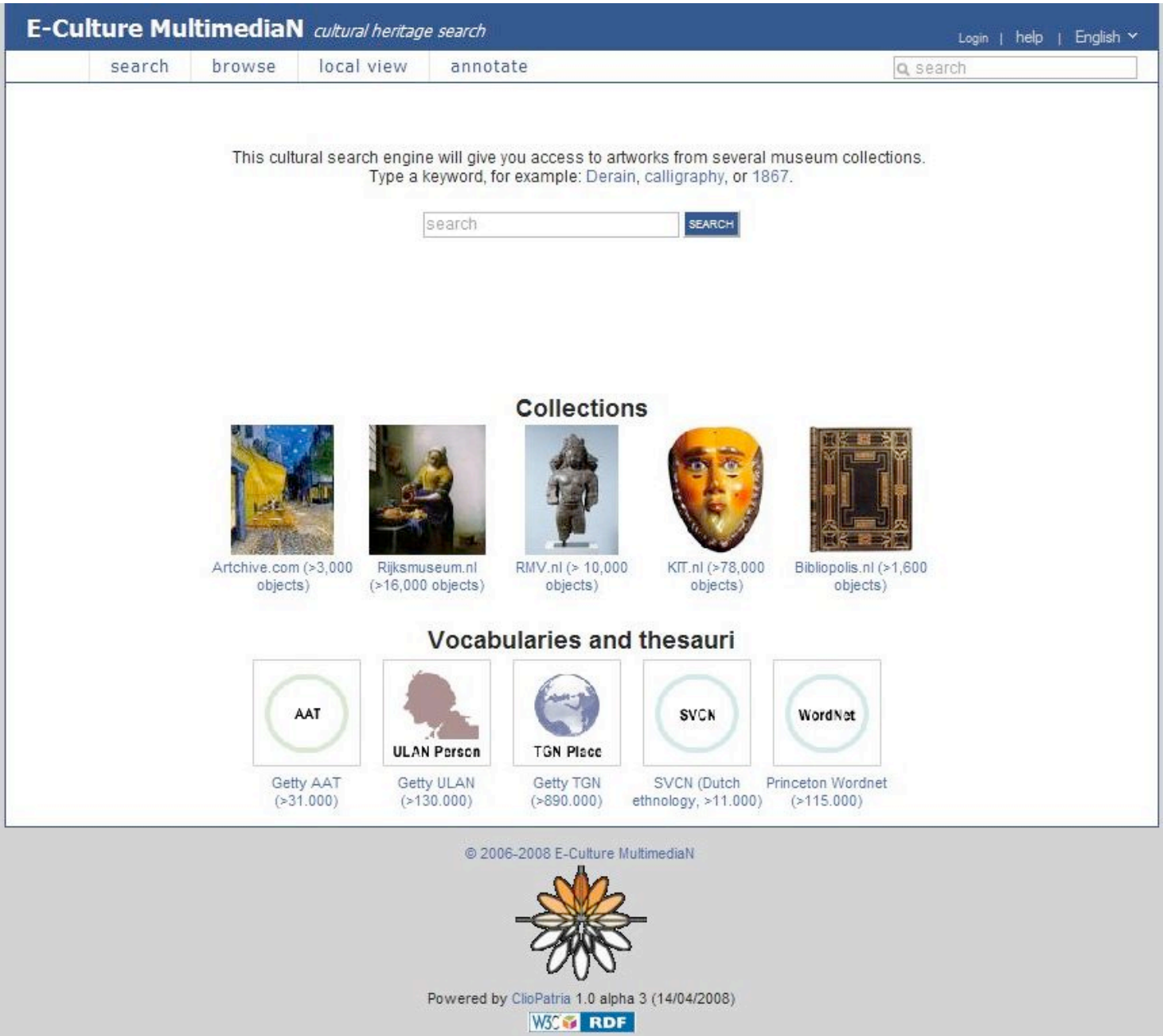
Who: Those interested in cultural heritage

Why: Exploring artefacts available in repository

What: Search combined collections

How: autocompletion to suggest topics, organise results





http://e-culture.multimedian.nl/demo/session/search

Use of linked data in E-Culture demo

- Query construction
  - auto-completion uses strings found in “data” and “concepts”
  - suggestions are grouped and ordered using links among items
- Result set
  - uses empirical balance between “closeness” to search string and non-intuitive path
- Result presentation
  - uses grouping of result set to show breadth of results
  - uses no particular ordering within each group

Subject Annotation

The execution of Johan van Oldenbarnevelt

RP-P-OB-77.320

A print that depicts the beheading of

Who Historical persons

person or organization

Oldenbarnevelt, Johan van x

What Iconclass/WordNet (en), events (nl)

(mythological) concept, object or event

beheading x

Where Name of place or region

geographical place

Den Haag x

When Date, year or period

enter date

Who: Professional annotators

Why: Subject matter annotation of 700.000 prints

What: Search in multiple thesauri for annotation terms

How: Autocompletion on who/what/where/when

http://e-culture.multimedian.nl/pk/annotate

Use of linked data in annotation task

- Query construction
  - auto-completion compares string in query with terms in thesauri
- Result set (the set of terms used to construct the menu)
  - terms that contain the string
- Result presentation (in the selection menu)
  - uses grouping of results depending on entry field
  - ordering also dependent on entry field
  - presentation of additional information differs per thesaurus and annotation field

Michiel Hildebrand, Jacco van Ossenbruggen, Lynda Hardman and Geertje Jacobs. Supporting subject matter annotation using heterogeneous thesauri, a user study in web data reuse. Technical Report INS-E0902, CWI, February 2009. http://ftp.cwi.nl/CWIreports/INS/INS-E0902.pdf

Conclusion

- Build specific application
- Determine information need
- Select data sources for task
- Ensure access to provenance information without being intrusive
  - remember hyperlink markers 20 years ago?
- Investigating re-usable interface components
  - autocompletion