ArtFinder: A Faceted Browser for Cross-Cultural Art Discovery

Matt Thompson

University of Bath, NII Tokyo, Sysemia Ltd Bristol

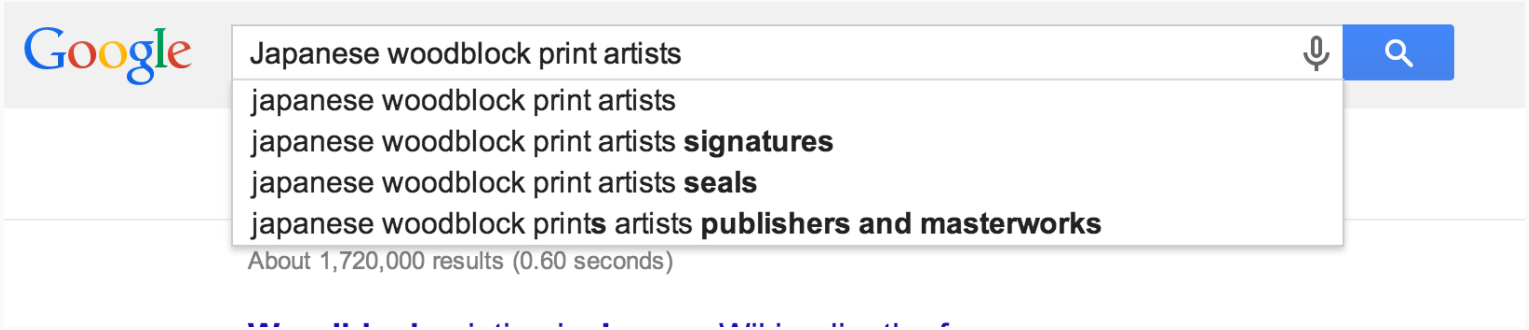
The problem

- How to explore art from another culture?
- Generally, you know what you like
- But what if the domain is totally unknown?

Possible solutions

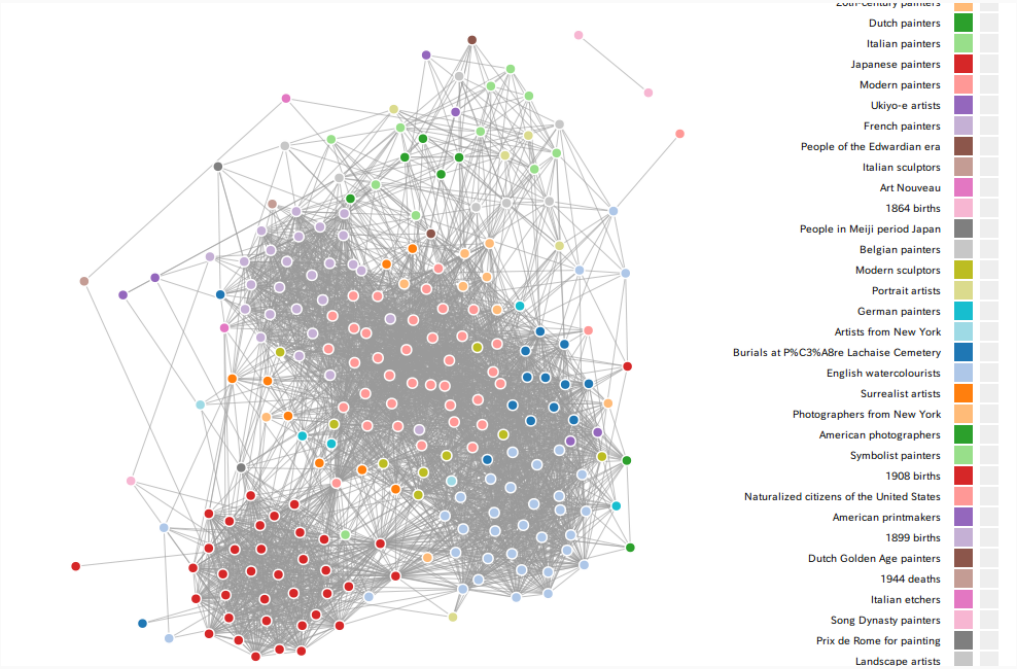
- Search
- Visualisation
- Categorisation and browsing

Search



- This is the standard approach
- What keywords would you use?
- What if the genres are words in a foreign language?

Visualisation



- We created ‘ArtViz’, a prototype visualisation
- Users can select genres of interest and ‘see’ similar artists
- Could be used for genres, eras, etc

Browsing the artists

- One problem: browsing artists from a *list* of tags is not ideal
- Ideally, the tags would be in some kind of hierarchical taxonomy
- How can we make this?

Faceted search

- First devised by Ranganathan [3] for library classification
- Sorts items into distinct, mutually exclusive facets
- Film example: ‘year’, ‘cast’, ‘genre’
- Many examples of faceted browsers, such as FLAMENCO browser for buildings [4]

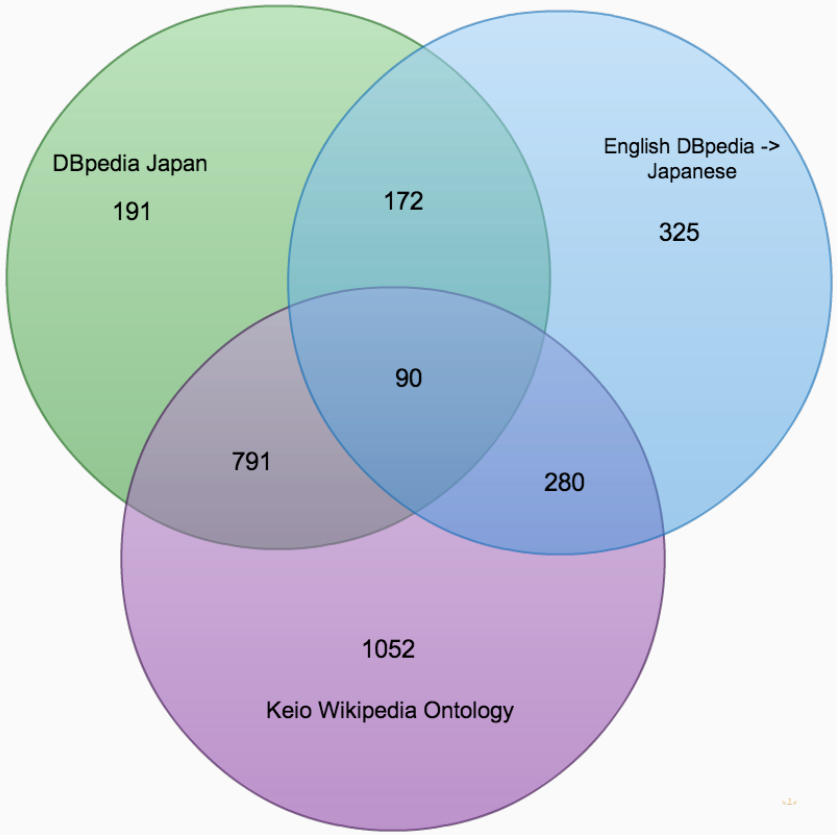
Facets

- Facets: *era, location, genre, media*
- Determined based on Ranganathan's [3] guidelines
- (Temporal, Spatial, Personal, Material)

Linked Open Data

- LODAC (`http://lod.ac`): Japanese Artists & Museums
- English DBpedia (Wikipedia articles)
- DBpedia Japan
- University of Keio's Japanese Wikipedia ontology

Datasets



- Cross-referenced LODAC with other ontologies
- Overlap (minus blanks/duplicates) of 893 usable artists

Extraction and translation of tags

- SPARQL queries sent to 893 artists
- Tags (genre, era, etc) taken from all ontologies
- Translated into English using Google Translate API

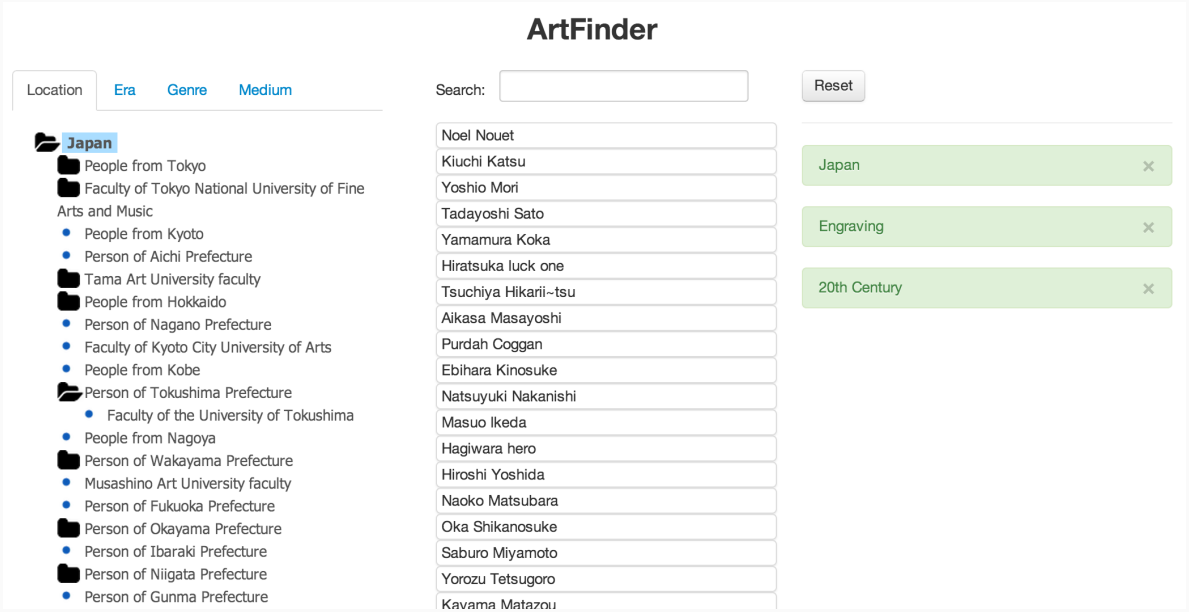
Hierarchy generation

$$P(x|y) = 1, P(y|x) < 1$$

$$D_x, D_y > 4$$

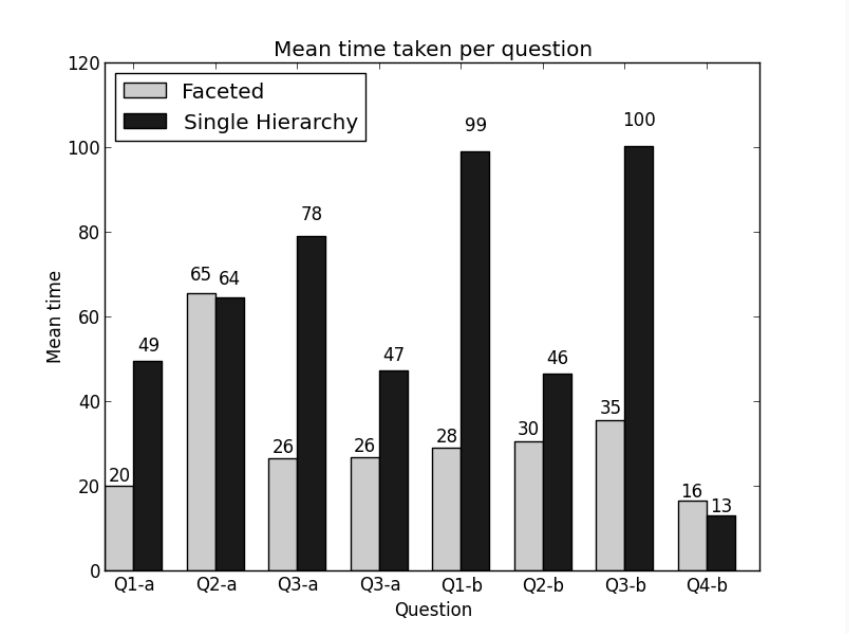
- Used Sanderson and Croft's subsumption approach [1]
- Relaxed from $P(x|y) = 1$ to $P(x|y) \geq 0.8$ to give better results
- Added D_x and D_y (number of documents in which tag occurs) from Schmitz et al [2]
- (Tag must appear > 4 times to be included in the hierarchy)

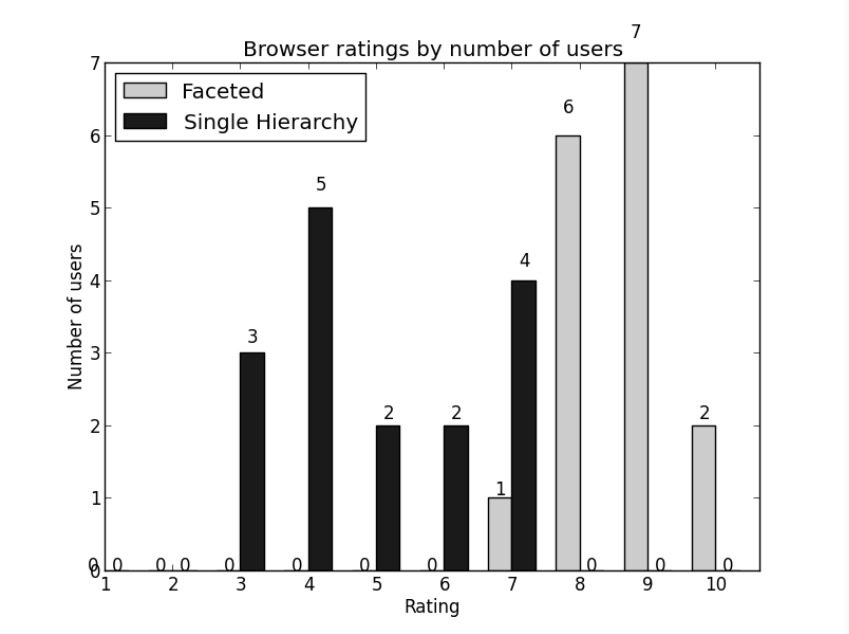
Browsing interface



- `http://is.gd/artfinder`
- Javascript, Angular.js, JSON
- Multiple tags can be selected from left

Preliminary user study





(a) Mean time taken for each interface, **(b)** Ratings for each interface, by user by question

- Faceted vs non-faceted browsers tested
- Given tasks to complete and timed

Conclusions, future work

Conclusions:

- Vast majority preferred faceted browser
- Faceted was preferred for both task completion and free exploration

Future work:

- Compare faceted browser with graph visualisation
- Genre/medium distinction unclear
- Tag names: France vs French. Fuzzy search?
- Automatically determine facets

References

Sanderson, Mark and Croft, Bruce
Deriving concept hierarchies from text.
Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pg. 206-213, 1999.

Schmitz, Patrick
Inducing ontology from flickr tags
Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland, 2006.

Ranganathan, Shiyali Ramamrita
Prolegomena to library classification
1967.

Hearst, Marti and Elliott, Ame and English, Jennifer and Sinha, Rashmi and Swearingen, Kirsten and Yee, Ka-Ping
Finding the flow in web site search
Communications of the ACM, 45(9):42-49, 2002.