Relational Semantic Search: Searching Social Paths on the Semantic Web

Jussi Kurki and Eero Hyvönen

Semantic Computing Research Group (SeCo)
Helsinki University of Technology (TKK) and University of Helsinki
first.last@tkk.fi, http://www.seco.tkk.fi/

Abstract. This paper presents a system searching for semantic relations between web resources, in our case historical persons. The system is based on the ULAN metadata of some 100,000 persons and organizations.

1 Relational Search

Semantic association identification has been studied in national security applications [1]. We have applied this notion[1], to be called *relational semantic search*, to cultural domain. The idea is to make it possible for the end-user to formulate queries such as "How is X related to Y" by selecting the end-point resources. The result is a set of semantic connection paths between X and Y. For example, in figure 1 the user has specified two historical persons, the Finnish artist Akseli Gallen-Kallela (1865–1931) and the French emperor Napoleon I (1769–1821) in a prototype[2] of the portal CultureSampo [2]. The system has discovered an association between the persons based on a chain of eight patronOf, teacherOf, knows, and studentOf relations.

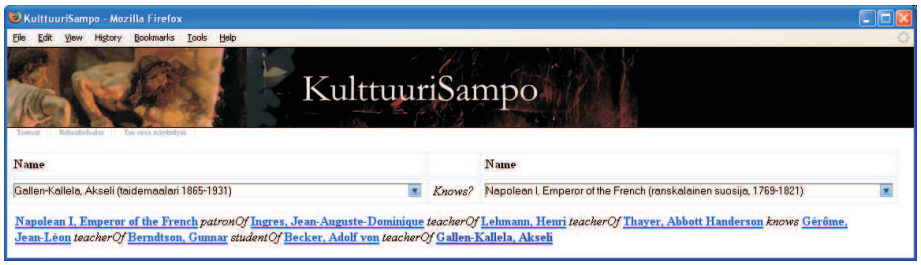

Fig. 1. Relational search in CultureSampo [2] using the ULAN vocabulary..

[1] A related work is http://e-culture.multimedian.nl/demo/search
[2] The relational search demo is available at http://demo.seco.tkk.fi/toimo/ff.html

2 Actor Ontology

Our system includes extensive information about artists based on the Union List of Artist Names (ULAN)[3] vocabulary. ULAN consists of over 100,000 individuals and corporate bodies of art historical signifigance, including comprehensive information about relationships between actors.

The model for our actor ontology is based on FOAF[4], Relationship[5] and BIO[6] vocabularies. Additional properties were added for roles and nationalities described in ULAN. The key property for our application is foaf:knows and its some 50 subproperties, such as childOf, parentOf, studentOf, teacherOf, friendOf.

ULAN data was converted to FOAF using XSLT-transformations. A memory-based graph was built from the data and the graph was stored as adjacency list. (The implementation was done in Java.) To minimize memory consumption, graph node has only minimal set of fields: an id and a list of children. At this point, all relationships are basically reduced to "knows" and all data is reduced to URI. Serialized to disk, the whole graph takes about 10MB of memory.

Relational search is done breath-first and even the longest paths (about 12 steps) can be found in less than half a second. This is explained partly by the structure of ULAN data. The graph has a strongly connected component of about 12000 actors containing central artists, such as Picasso and Donatello. At the same time, thousands of others, especially contemporary artists, don't have many connection in the underlying RDF graph.

Acknowledgements

This research is part of the National Finnish Ontology Project (FinnONTO) 2003-2007[7], funded mainly by The National Technology Agency (Tekes) and a consortium of 36 companies and public organisations.

References

1. Sheth, A., Aleman-Meza, B., Arpinar, I.B., Bertram, C., Warke, Y., Ramakrishnan, C., Halaschek, C., Anyanwu, K., Avant, D., Arpinar, F.S., Kochut, K.: Semantic association identification and knowledge discovery for national security applications. Journal of Database Management on Database Technology **16**(1) (2005) 33–53
2. Hyvönen, E., Ruotsalo, T., Häggström, T., Salminen, M., Junnila, M., Haaramo, M.V.M., Kauppinen, T., Mäkelä, E., Viljanen, K.: CultureSampo—Finnish culture on the semantic web. The vision and first results. In: Semantic Web at Work—Proceedings of STeP 2006. To appear in: Klaus Robering (Ed.), Information Technology for the Virtual Museum. LIT Verlag, 2007. (2006)

[3] http://www.getty.edu/research/conducting_research/vocabularies/ulan/
[4] http://xmlns.com/foaf/spec/
[5] http://vocab.org/relationship/
[6] http://vocab.org/bio/0.1/
[7] http://www.seco.tkk.fi/projects/finnonto/

Demonstration

The relational search component of CultureSampo will be demonstrated at the conference on-line: http://demo.seco.tkk.fi/toimo/ff.html