Analyzing and Visualizing Prosopographical Linked Data Based on Short Biographies

Petri Leskinen[1], Eero Hyvönen[1,2], and Jouni Tuominen[1,2]

[1] Semantic Computing Research Group (SeCo), Aalto University, Finland, and
[2] HELDIG – Helsinki Centre for Digital Humanities, University of Helsinki, Finland
`http://seco.cs.aalto.fi, http://heldig.fi`
`firstname.lastname@aalto.fi`

Type of submission: original unpublished work

In our earlier paper [1] we presented an application case study where data from a printed collection of some 10,0000 short biographies of high school alumni was extracted and transformed into Linked Open Data, enriched by data linking to 10 external data sources, and published in a SPARQL endpoint. On top of the data service, a semantic faceted search engine and browser was developed for searching and filtering persons/biographies. This paper extends this work by showing how faceted search can be utilized as a basis for prosopographical data analysis and research: a new application is presented where various data visualization tools using Google Charts have been integrated with the SPARQL endpoint allowing the end user to filter out subsets of persons/biographies, and then to study them. In addition to providing statistical analyses of person groups, an interesting use case identified here is to compare visualizations based on different subgroups, e.g., famous people with entries in related datasets and those not included there. In this way it is possible, for example, to determine what education, profession, or employer will most likely lead to having an entry in the National Biography of Finland or Wikipedia. The data service is available at the Linked Data Finland platform [2], including some 892,000 triples about 131,000 resources. The extended application [3] is now in use on the Semantic Web.

[1] Eero Hyvönen, Petri Leskinen, Erkki Heino, Jouni Tuominen and Laura Sirola: Reassembling and Enriching the Life Stories in Printed Biographical Registers: Norssi High School Alumni on the Semantic Web. Language, Technology and Knowledge. First International Conference, LDK 2017, Galway, Ireland, June 19-20, 2017, Proceedings. Springer-Verlag,

[2] http://ldf.fi/norssit/sparql

[3] http://www.norssit.fi/semweb