Semantic Web 0 (0) 1
IOS Press

Using the Semantic Web in Digital Humanities: Shift from Data Publishing to Data-analysis and Serendipitous Knowledge Discovery

Eero Hyvönen
University of Helsinki, Helsinki Centre for Digital Humanities (HELDIG), Finland and
Aalto University, Department of Computer Science, Finland
Semantic Computing Research Group (SeCo) (http://seco.cs.aalto.fi)
E-mail: eero.hyvonen@aalto.fi

Editors: Pascal Hitzler, Kansas State University, Manhattan, KS, USA; Krzysztof Janowicz, University of California, Santa Barbara, USA
Solicited reviews: Rafael Goncalves, Stanford University, CA, USA; Peter Haase, metaphacts GmbH, Walldorf, Germany; One anonymous reviewer

Abstract. This paper discusses a shift of focus in research on Cultural Heritage semantic portals, based on Linked Data, and envisions and proposes new directions of research. Three generations of portals are identified: Ten years ago the research focus in semantic portal development was on data harmonization, aggregation, search, and browsing ("first generation systems"). At the moment, the rise of Digital Humanities research has started to shift the focus to providing the user with integrated tools for solving research problems in interactive ways ("second generation systems"). This paper envisions and argues that the next step ahead to "third generation systems" is based on Artificial Intelligence: future portals not only provide tools for the human to solve problems but are used for finding research problems in the first place, for addressing them, and even for solving them automatically under the constraints set by the human researcher. Such systems should preferably be able to explain their reasoning, which is an important aspect in the source critical humanities research tradition. The second and third generation systems set new challenges for both computer scientists and humanities researchers.

Keywords: Digital Humanities, Linked Data, Semantic portals, Data analysis, Knowledge discovery

1. Introduction

Cultural Heritage (CH) has become a most active area of application of Linked Data and Semantic Web (SW) technologies [1]. Large amounts of CH content and metadata about it are available openly for research and public use based on collections in museums, libraries, archives, and media organizations. For example, data has been aggregated in large national and international repositories, web services, and portals such as Europeana[1] and Digital Public Library of America[2], and forms a substantial part of DBpedia[3] and Wikidata[4].

The availability of Big Data has boosted the rapidly emerging new research area of Digital Humanities (DH) [2, 3] where computational methods are developed and applied to solving problems in humanities

[1] http://europeana.eu
[2] https://dp.la/
[3] http://dbpedia.org
[4] http://wikidata.org

and social sciences. In this context Big Data means data that is too big or complex to be analyzed manually by close reading [4].

From a SW research point of view, CH data provide interesting challenges for DH research: First, the data is syntactically heterogeneous (text, images, sound, videos, and structured data in different formats, such as XML, JSON, CSV, and RDF) and written in different languages. Second, the data is semantically rich covering all aspects of life in different times and places. Third, the data are often incomplete, imprecise, uncertain, or fuzzy due to the nature of history. Fourth, the data is interlinked across different data sources, distributed in different countries and databases. Helping the humanities researcher to deal with such data in semantically complex problems addressed in humanities sets for computer scientists interesting methodological problems.

This paper analyses and discusses this line of research and development at the crossroads of Semantic Web research, humanities, and social sciences, from the early days of the Semantic Web to next steps in the future. Three conceptual generations of semantic portals on the Semantic Web are first identified. After this the ideas are made more concrete by an example case study system exhibiting features of the three generations.

2. First Generation: Portals for Search and Browsing

Due to the challenges in CH data, SW research in CH has been initially focused on issues related to syntactic and semantic interoperability and data aggregation. A great deal of work has been devoted in developing metadata standards and data models for harmonizing data, including application agnostic W3C standards[5] (RDF, OWL, SKOS, etc.), document centric models, such as Dublin Core and its dumb down principle[6], and event-centric models for data harmonization on a more fundamental level, such as CIDOC CRM[7] [5] for museums and its extensions[8], and FRBRoo [6] and IFLA Library Reference Model (LRM)[9] in libraries. In document-centric metadata models the idea is to agree upon a shared way of describing the properties of documents, and how different models can be mapped on each other for interoperability. The event-centric approach focuses on developing more fundamental ontological models of the real world onto which different data and metadata can be transformed for interoperability. Once the data is harmonized in one way or another, it can be published in a SPARQL endpoint, and semantic portals can be created on top of it via APIs.

Both document-centric and event-centric approaches have been successful. Dublin Core and its extensions have become the metadata norm for representing documents on the Web, and a lot of use cases and applications of CIDOC CRM[10] and other event-centric systems have been published.

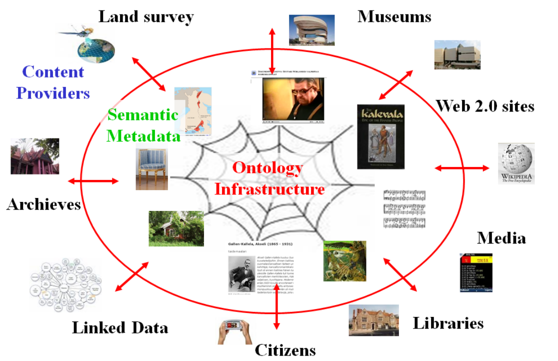

Fig. 1. A model for distributed Linked Data publishing. The data publishers around the circle, i.e., a joint publishing system, provide data using the vocabularies of a shared ontology infrastructure in the middle. The data are automatically interlinked and enrich each other.

The ideas of the Semantic Web and Linked Data can be applied to address the problems of semantic data interoperability and distributed content creation at the same time, as depicted in Fig. 1. Here the publication system is illustrated by a circle. A shared semantic ontology infrastructure is situated in the middle. It includes shared domain ontologies, modeled using SW standards. If content providers outside of the circle provide the system with metadata about CH based on the same ontologies, the data are automatically linked through shared URIs, enrich each other, and form a joint knowledge graph.

For example, if metadata about a painting created by Picasso comes from an art museum, it can be enriched

[5] http://www.w3.org/standards/semanticweb
[6] https://www.dublincore.org/
[7] http://cidoc-crm.org
[8] http://www.cidoc-crm.org/collaborations
[9] https://www.ifla.org/publications/node/11412
[10] http://www.cidoc-crm.org/useCasesPage

by data links to, e.g., biographies from Wikipedia and other sources, photos taken of Picasso, information about his wives, books in a library describing his works of art, related exhibitions open in museums, and so on. At the same time, the contents of any organization in the portal having Picasso-related material get enriched by the metadata of the new artwork entered in the system. This is a win-win business model for everybody to join such a system; collaboration pays off.

Combining the infrastructure with the idea of decoupling the data services for machines from the applications for the human user creates a model for building collaborative Semantic Web applications. This model has been developed and tested in practice, e.g., in the “Sampo” series of semantic portals[11] [7, 8]. The idea of collaborative content creation using Linked Data has been developed also in other settings, e.g., in ResearchSpace[12].

The main use case in CH portals has been providing the user with enhanced information retrieval (IR) facilities [9], such as faceted search [10], semantic search, entity search, and semantic recommendation systems [11] for exploring the data in intelligent ways. Such CH search and browsing systems based on harmonized aggregated linked data will be called *first generation systems*.

3. Second Generation: Portals with Tools for Distant Reading

As more and more harmonized aggregated linked datasets are available, the time has come to take a next step forward to *second generation* of CH semantic portals. The novelty of such systems is to provide the user with tools for solving Digital Humanities (DH) research problems, not only tools for searching and browsing the data. For example, the researcher may be interested in finding out, how historical persons, ships, or manuscripts have been moving around geographically, what topics have appeared and when in parliamentary discussions, newspapers, or other corpora, what kind of social networks or correspondences there have been between members of a society, and so on. In DH, a key goal is to use computational methods for solving humanities and social science problems using large datasets that have become available. A variety of technologies have been developed and applied for such tasks, such as sentiment analysis [12], topic modeling [13], network analysis [14, 15], and visualizations [16] in addition to traditional and novel statistical methods, such as word embeddings and neural networks [17–19].

Many of the methods and tools above are domain independent, and there are a lot of software packages available for using them, such as Gephi[13], R [20], and various Python libraries[14]. However, each of them have their own input formats and user interfaces. Furthermore, visualizations are crafted case by case; tools for formulating, adjusting, and comparing them in generalizable ways would be helpful for the user. A major problem here is that using the tools typically requires technical expertise and skills not common among the humanities researchers. Furthermore, the tools usually do not support Linked Data formats and data services, and there is the burden of transforming and transporting linked data into formats required by the different data analysis tools. A challenge therefore is how to create the tools in a generalizable way so that the end user can adapt them for her own particular research problems.

At the moment, many portals include tools but they are mostly aimed for visualizing and exploring the data. Showing data on maps and timelines are common examples of this. The same applies to some systems for network analysis, such as Six Degrees of Francis Bacon[15], where one can search for a person or a group whose networks are then shown for exploration using interactive graphs. It is also possible to show the group on a timeline and, e.g., filter the connections in the network based on parameters. To move on to second generation systems with a clearer focus on data analytic tooling one could, e.g., compute various connectivity parameters and statistics of the networks, such as most connected nodes, hubs, and connections in the data. Such tool-oriented systems are largely still missing in semantic portals; data analysis in Digital Humanities is usually done by downloading data and by transforming it locally to be used in specific off-line tools. Integration of data analytic tools with online semantic portals is a promising future direction of work [21, 22].

[11] https://www.europenowjournal.org/2019/09/09/linked-data-in-use-sampo-portals-on-the-semantic-web/
[12] https://www.researchspace.org/
[13] https://gephi.org/
[14] https://bigdata-madesimple.com/top-20-python-libraries-for-data-science/
[15] http://www.sixdegreesoffrancisbacon.com/

4. Third Generation: Portals for Serendipitous Knowledge Discovery

Current DH systems have focused on semantic data aggregation, enrichment, validation, search, exploration, visualization, and in some cases even data analysis. The idea has been to search and present the data to the DH researcher using statistical charts, maps, timelines, graphs, and other means so that the researcher can more easily analyze the data related to her/his research problem. What is still largely missing in the DH methodology and tools is the next conceptual level of Artificial Intelligence where the DH tool is able not only to present the data to the human researcher in useful ways but also to 1) find, address, or solve the DH research problems *automatically by itself* and 2) also explain its reasoning or solution to the researcher. This is a grand challenge for research in the future.

To address this challenge on has to study serendipitous[16] knowledge discovery (KD) [23, 24] in the context of historical Cultural Heritage Big Data. Another direction of research to draw ideas from is Computational Creativity in Artificial Intelligence [25]. Serendipitous knowledge discovery is one of original promises of the Semantic Web [26]. However, there is surprisingly little research about it. A reason for this may be shortage of high quality densely interlinked datasets needed for studying serendipity. Also the notion of serendipity is conceptually complicated. Better understanding of the notion of serendipity and how insightful knowledge discovery can be implemented and utilized is needed. This could lead to new insights of scientific discovery in humanities and to a paradigm change where the role of the computer is changing from a passive tool to a proactive intelligent agent.

For this challenge the research agenda for the future should seek answers to, e.g., the following fundamental research questions:

1. *How can one formalize the notion of serendipity in terms of 'interestingness' [27] in a generalizable way?* It does not make sense to hard code serendipity in a system using specific ad hoc rules, otherwise reasoning would not be serendipitous.
2. *How can serendipitous phenomena and their explanations be extracted from the data?*
3. *How can the notion of serendipity (1) and the methods for discovering it (2) be used in practice for finding, addressing, and solving humanities research problems?*
4. *How can semantically rich-enough linked datasets for (1–3) be created, based on combining both structured and non-structured data?* An important research topic here is Natural Language Understanding, since the primary data is typically available in textual forms.

In previous sections, semantic portals have been categorized conceptually into three generations. However, in practise the later generation systems have to address the challenges of the former generations, too: a requisite for both second and third generation systems is availability of harmonized linked data, as in first generation systems, and third generation systems also focus on tools in a way similar to second generation systems.

In order to make the ideas presented above more concrete by an example, a semantic portal, BiographySampo, is presented next. This system was created with the goal of making a paradigm shift in its field from state-of-the art first generation systems to a second generation systems. However, the system also includes a third generation tool for serendipitous knowledge discovery.

5. A Case Study: BiographySampo – Biographies on the Semantic Web

Biography is a research area in humanities that studies life stories of particular people of significance, with the aim of getting a better understanding of their personality and actions, e.g., to understand their motives [28]. An important resource in this research field are biographical dictionaries [29] that may contain tens of thousands of short biographies of historical persons of importance[17]. Traditionally, such dictionaries have been published as printed book series but nowadays major biographical dictionaries have opened their editions on the Web with search engines for finding and (close) reading biographies of interest.

[16] Serendipity means 'happy accident' or 'pleasant surprise', even 'fortunate mistake'.

[17] On-line national biographical collections include, e.g., USA's American National Biography [30], Germany's Neue Deutsche Biographie [31], Biography Portal of the Netherlands [32], Dictionary of Swedish National Biography [33], and National Biography of Finland [34].

In BiographySampo[18] [22], linked data and natural language technology was used for creating a knowledge graph encompassing the data related to 13 100 biographies, including the National Biography of Finland [34]. The data was harmonized using an extension of CIDOC DRM and was linked to 16 external datasets for enriching the contents. The data was published in a SPARQL endpoint, and faceted search was implemented on top of the data service for finding biographies and exploring them by browsing. These features make BiographySampo a state-of-the-art first generation system.

In contrast to biography, the focus of *prosopography* research is to study life histories of groups of people in order to find out some kind of commonness or average in them [35]. For example, the research question may be to find out what happened to the students of a school in terms of social ranking and employment after their graduation. The prosopographical research method [35, p. 47] has two steps: First, a target group of entities in the data is selected that share desired characteristics for solving the research question at hand. Second, the target group is analyzed, and possibly compared with other groups, in order to solve the research question. The analysis may involve, e.g., creating pie charts, histograms or other statistics of the target group, mapping the target group geographically, network analysis, etc.

To support prosopography, a second generation CH application with tooling is needed. Filtering out the target group is not enough but tools and visualizations are needed for analyzing it, too. In developing BiographySampo, a major goal has been in providing the DH researchers with generic tools for data visualization and analysis. Moreover, the tools can be applied not only to one target group but also to two parallel groups in order to compare them. For example, Fig. 2 compares the life charts of Finnish generals and admirals in the Russian armed forces in 1809–1917 when Finland was an autonomous Grand Duchy within the Russian Empire (on the left) with the members of the Finnish clergy (1800–1920) (on the right). With a few selections from the facets the user can filter out the two target groups and see that, for some reason, quite a few officers moved to Southern Europe when they retired (like retirees today) while the Lutheran ministers tended to stay in Finland.

[18]The portal is online at http://biografiasampo.fi and has had tens of thousands of users

In the same way, the statistical application perspective in the system includes histograms showing various numeric value distributions of the members of the target groups, e.g., their ages, number of spouses and children, and pie charts visualizing proportional distributions of professions, societal domains, and working organizations. There is also a network perspective based on the idea of visualizing and studying networks among target groups filtered out using facets. The networks are based on the reference links between the biographies, either handmade or based on automatically detected mentions. The depth of the networks can be controlled by limiting the number of links, and coloring of the nodes can be based on the gender or societal domain of the person (e.g., military, medical, business, music, etc.).

The biographies can also be analyzed as a collection of artefacts by using linguistic analysis. For example, it turns out that the biographies of female Members of the Parliament (MP) frequently contain words ”family” and ”child”, but these words are seldom used in the biographies of male MPs. The analyses are based on a linguistic knowledge graph of the texts.

These tools and functionalities make BiographySampo a second generation system. To study and explore the possibilities and challenges of third generation systems, yet another application perspective was created in BiographySampo for finding interesting serendipitous connections in the biographical knowledge graph. This application idea is related to relational search [36, 37]. In our case a new knowledge-based approach was developed to find out in what ways (groups of) people are related to places and areas. Such connections can reveal hidden indirect relations that are new and surprising to the user. This method, described in more detail in [38], rules out non-sense relations effectively and is able to create natural language explanations for the connections.

The question to be solved is formulated by making selections on facets about people, professions, places, and generic relation types. For example, the question ”How are Finnish artists related to Italy?” is solved by selecting ”Italy” from the place facet and ”artist” from the profession facet. The results include connections between people and places constrained by the facet selections, e.g., that ”Elin Danielson-Gambogi received in 1899 the Florence City Art Award” and ”Robert Ekman created in 1844 the painting ’Landscape in Subiaco’ depicting a place in Italy”. Finding out hidden “new” semantic associations and their explanations like these in a large knowledge graph (over

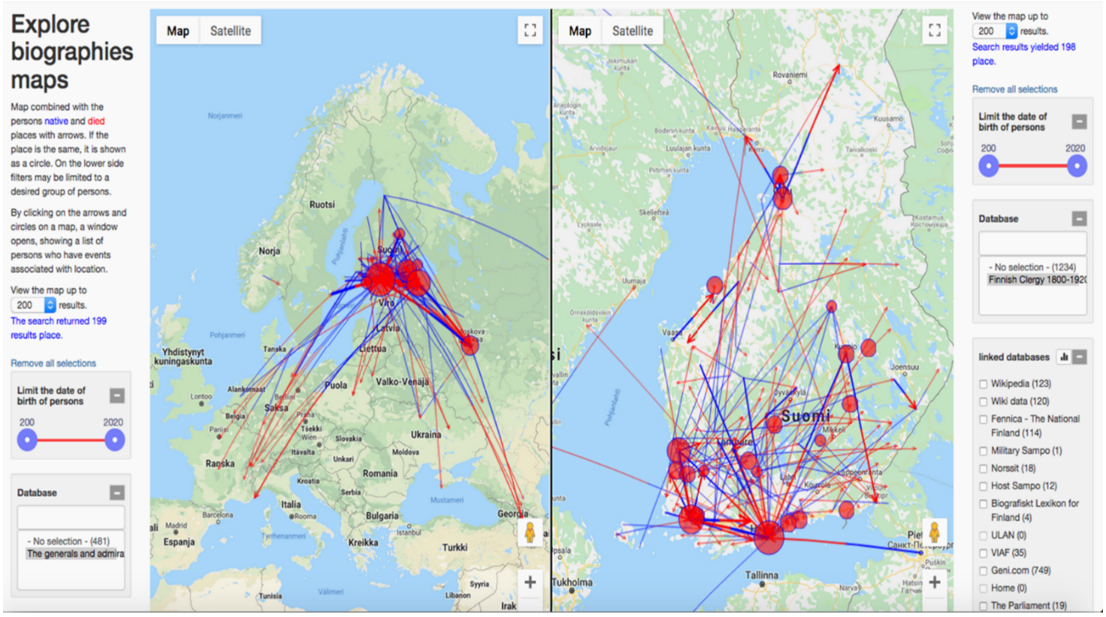

Fig. 2. Comparing the life charts of two target groups, admirals and generals (left) and clergy (right) of the historical Grand Duchy of Finland (1809–1917).

10 million triples), created using the model of Fig.1, can arguably be considered serendipitous knowledge discovery. This makes BiographySampo an example of a third generation semantic portal. Knowledge discovery in this application is performed by transforming the knowledge graph into instances of serendipitous connections and their explanations in a preprocessing phase using rule-based reasoning. After this, relational search can be reduced into faceted search on the connection instances.

6. Conclusions

This paper discussed how focus in developing semantic portals for Cultural Heritage has been evolving during the last 10 years, and proposes and envisions next steps ahead. A three generation model was presented for characterising the process: The first generation systems provided the end user with search and browsing facilities on top of a data service of harmonized linked data (SPARQL endpoint). The second generation systems provide the user also with data-analytic tools that help the Digital Humanities researcher in addressing and solving research problems. In the envisioned third generation systems a step on a new conceptual level towards Artificial Intelligence is taken: the role of the portal is not only to provide tools for the human researcher to use but also actively and automatically find interesting serendipitous patterns in the data and even solve problems by itself, preferably with explicit explanations. In addition to knowing that the meaning of life is “42”, as suggested by the computer in the novel *Hitchhiker’s Guide to the Galaxy* by Douglas Adams, we also need to know why so.

This shift of research focus from data publishing to data analysis and tooling and finally to Artificial Intelligence brings in novel research challenges in, e.g., knowledge extraction, data visualization, machine learning, knowledge discovery, and computational creativity. Interpreting the results of a tool typically requires a great deal of domain knowledge and understanding the underlying algorithms and the characteristics of the data, such as modeling principles used and completeness, uncertainty, and fuzziness of the data. Using advanced computational tools in Digital Humanities raises the demand for source criticism on a new, higher level.

References

[1] E. Hyvönen, *Publishing and using cultural heritage linked data on the semantic web*, Morgan & Claypool, Palo Alto, CA, 2012, https://doi.org/10.2200/S00452ED1V01Y201210WBE003.

[2] W. McCarty, *Humanities Computing*, Palgrave, London, 2005.
[3] E. Gardiner and R.G. Musto, *The Digital Humanities: A Primer for Students and Scholars*, Cambridge University Press, New York, NY, USA, 2015, https://doi.org/10.1017/CBO9781139003865.
[4] K. Shultz, What Is Distant Reading?, *New York Times* (June, 24, 2011).
[5] M. Doerr, The CIDOC CRM—an Ontological Approach to Semantic Interoperability of Metadata, *AI Magazine* **24**(3) (2003), 75–92.
[6] C. Bekiari, M. Doerr, P.L. Bœuf and P. Riva (eds), *Definition of FRBRoo. A Conceptual Model for Bibliographic Information in Object-Oriented Formalism*, International Federation of Library Associations and Institutions (IFLA), 2015, https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf.
[7] E. Hyvönen, Cultural Heritage Linked Data on the Semantic Web: Three Case Studies Using the Sampo Model, in: *VIII Encounter of Documentation Centres of Contemporary Art: Open Linked Data and Integral Management of Information in Cultural Centres Artium, Vitoria-Gasteiz, Spain, October 19-20, 2016*, 2016.
[8] E. Hyvönen, "Sampo" Model and Semantic Portals for Digital Humanities on the Semantic Web, in: *Digital Humanities in Nordic Countries (DHN 2020), Proceedings, accepted*, 2020, https://seco.cs.aalto.fi/publications/2020/hyvonen-sampos-dhn-2020.pdf.
[9] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval (2nd Edition)*, Addison-Wesley, New York, 2011.
[10] D. Tunkelang, Faceted search, *Synthesis Lectures on Information Concepts, Retrieval, and Services* **1**(1) (2009), 1–80, https://doi.org/10.2200/S00190ED1V01Y200904ICR005.
[11] D. Jannach, M. Zanker, A. Felfernig and G. Friedrich, *Recommender Systems. An introduction*, Cambridge University Press, Cambridge, UK, 2011.
[12] B. Liu, *Sentiment Analysis and Opinion Mining*, Morgan & Claypool, Palo Alto, CA, 2012, https://doi.org/10.2200/S00416ED1V01Y201204HLT016.
[13] M.R. Brett, Topic Modeling: A Basic Introduction, *Journal of Digital Humanities* **2**(1) (2012). http://journalofdigitalhumanities.org/2-1/topic-modeling-a-basic-introduction-by-megan-r-brett/.
[14] E. Otte and R. Rousseau, Social network analysis: a powerful strategy, also for the information sciences, *Journal of information Science* **28**(6) (2002), 441–453.
[15] C.N. Warren, D. Shore, J. Otis, L. Wang, M. Finegold and C. Shalizi, Six Degrees of Francis Bacon: A Statistical Method for Reconstructing Large Historical Social Networks., *DHQ: Digital Humanities Quarterly* **10**(3) (2016).
[16] A.-S. Dadzie and M. Rowe, Approaches to Visualising Linked Data: A Survey, *Semantic Web – Interoperability, Usability, Applicability* **1**(1–2) (2011).
[17] M. Chen, Efficient Vector Representation for Documents Through Corruption, in: *5th International Conference on Learning Representations*, OpenReview.net, 2017.
[18] Q. Le and T. Mikolov, Distributed representations of sentences and documents, in: *International Conference on Machine Learning*, 2014, pp. 1188–1196.
[19] T. Mikolov, I. Sutskever, K. Chen, G. Corrado and J. Dean, Distributed Representations of Words and Phrases and their Compositionality, *arXiv* (2013). http://arxiv.org/abs/1310.4546.
[20] A. Field, J. Miles and Z. Field, *Discovering Statistics Using R*, SAGE Publications Inc., USA, 2015.
[21] P. Leskinen, G. Miyakita, M. Koho and E. Hyvönen, Combining Faceted Search with Data-analytic Visualizations on Top of a SPARQL Endpoint, in: *Proceedings of VOILA 2018, Monterey, California*, Vol. 2187, CEUR Workshop Proceedings, 2018.
[22] E. Hyvönen, P. Leskinen, M. Tamper, H. Rantala, E. Ikkala, J. Tuominen and K. Keravuori, BiographySampo - Publishing and Enriching Biographies on the Semantic Web for Digital Humanities Research, in: *The Semantic Web. 16th International Conference, ESWC 2019, Portorož, Slovenia, June 2–6, 2019, Proceedings*, Springer–Verlag, 2019, http://doi.org/10.1007/978-3-030-21348-0.
[23] O. Maimon and L. Rokach (eds), *The data mining and knowledge discovery handbook*, Springer–Verlag, 2005, http://doi.org/10.1007/978-0-387-09823-4.
[24] G. Piatetsky-Shapiro and W. Frawley (eds), *Knowledge discovery in databases*, The MIT Press, Cambridge, Massachusetts, 1991.
[25] M.A. Boden, Computer Models of Creativity, *AI Magazine* **30**(3) (2009), 23–34. http://www.aaai.org/ojs/index.php/aimagazine/article/view/2254.
[26] O. Lassila, Serendipitous Interoperability, in: *The Semantic Web Kick-Off in Finland – Vision, Technologies, Research, and Applications*, Helsinki Institute for Information Technology, 2002, HIIT Publications 2002-001.
[27] A. Silbershactz and A. Tuzhilin, On subjective measures on interestingness in knowledge discovery, in: *Proceedings of KDD-1995*, AAAI Press, 1995.
[28] B. Roberts, *Biographical Research*, Understanding social research, Open University Press, 2002. ISBN 9780335202867.
[29] T. Keith:, *Changing conceptions of National Biography*, Cambridge University Press, 2004.
[30] American National Biography, http://www.anb.org/aboutanb.html.
[31] Neue Deutsche Biographie, http://www.ndb.badw-muenchen.de/ndb_aufgaben_e.htm.
[32] Biography Portal of the Netherlands, http://www.biografischportaal.nl/en.
[33] Dictionary of Swedish National Biography, https://sok.riksarkivet.se/Sbl/Start.aspx?lang=en.
[34] National Biography of Finland, https://kansallisbiografia.fi/english.
[35] K. Verboven, M. Carlier and J. Dumolyn, A short manual to the art of prosopography, in: *Prosopography approaches and applications. A handbook*, Unit for Prosopographical Research (Linacre College), 2007, pp. 35–70, http://dx.doi.org/1854/8212.
[36] S. Lohmann, P. Heim, T. Stegemann and J. Ziegler, The RelFinder User Interface: Interactive Exploration of Relationships between Objects of Interest, in: *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI 2010)*, ACM, 2010, pp. 421–422.
[37] G. Tartari and A. Hogan, WiSP: Weighted Shortest Paths for RDF Graphs, in: *Proceedings of VOILA 2018*, CEUR Workshop Proc., Vol. 2187, 2018, pp. 37–52.
[38] E. Hyvönen and H. Rantala, Knowledge-based Relation Discovery in Cultural Heritage Knowledge Graphs, in: *Proceedings of the 4th Digital Humanities in the Nordic Countries*

Conference (DHN 2019), CEUR Workshop Proceedings, Vol- 2364, 2019.