A User Interface Model for Digital Humanities Research: Case BookSampo – Finnish Fiction Literature on the Semantic Web

Annastiina Ahola[1][0009−0008−6369−4712], Eero Hyvönen[1,2][0000−0003−1695−5840], and Heikki Rantala[1][0000−0002−4716−6564]

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

Abstract. This paper presents the implementation of a new user interface (UI) for the Finnish fiction literature BOOKSAMPO knowledge graph. The UI utilizes possibilities of semantic web technologies to provide the end-user an enhanced search and browsing experience through faceted search and data-analytical tools for Digital Humanities research.

Keywords: Digital Libraries, Linked Data, User Interfaces, Portals

1 Introduction

BOOKSAMPO[3] is a semantic portal containing information on virtually all Finnish fiction literature. The portal makes use of the BOOKSAMPO knowledge graph (KG) consisting of Linked Data (LD) descriptions of literary works and other objects related to those works, e.g., authors and covers, that have been systemically recorded for works in Finnish libraries since 1997 [6]. While the usage of LD for library collections is not new [1], the BOOKSAMPO KG is a uniquely rich fiction literature data set on an international scale that offers countless possibilities for data analysis and literary research.

The currently available BOOKSAMPO PORTAL was published in 2011. It is part of the series of *Sampo portals*[4] based on the "Sampo Model" for LD publishing [3]. Its user interface (UI) is built with Drupal and the portal offers a text-based search engine for searching and exploring the underlying KG. While the search engine makes use of links between records, the full potential of what LD could offers for digital libraries is not realized in the portal. As a part of a project revisiting the BOOKSAMPO KG in 2022, a new semantic UI based on integrated data analytic tools and faceted search was developed for the BOOKSAMPO KG. This paper presents this new UI developed using the Sampo-UI Framework[5] [5] based on the previously mentioned Sampo Model.

---

[3] See project research homepage at: https://seco.cs.aalto.fi/applications/kirjasampo/
[4] See the homepage of Sampo portals at: https://seco.cs.aalto.fi/applications/sampo/
[5] See code and documentation at: https://github.com/SemanticComputing/sampo-ui

2 BookSampo Knowledge Graph

The original BOOKSAMPO PORTAL[6] deployed in 2011 was developed as a part of the national FinnONTO research initiative (2003–2012) [2]. The portal has then been maintained by the Finnish Public Libraries[7] and had more than 1.1 million distinct visitors in 2021[8]. The original data came from legacy library databases that were harmonized and transformed into RDF format. After this, the librarians have maintained and enriched the data [7] by themselves. The original transformed dataset contained more than 3 million triples with around 90 000 new triples being added monthly. Today the KG consists of nearly 9 million triples with over 210 000 abstract literary works and nearly 220 000 publications.

3 Using the New User Interface for the BookSampo KG

Based on the Sampo-UI framework, the landing page of the portal contains *application perspectives* through which instances of the major classes of the underlying KG can be searched. In this case, there are five perspectives: Three of them deal directly with literary works—*novels*, *nonfiction books*, and *publications*—and the other two deal with *authors* and *covers*. The novels and nonfiction books perspectives contain information about those books on the *abstract work level* while the publications perspective has information on all works on the *physical work level* following the conceptual distinction made in the original data [7].

Clicking on an application perspective card opens up a faceted search view of the perspective. The user can then use the facets to filter the results to match specific criteria or explore the result set as a whole. The perspectives offer various tabs for visualizing the result set in different forms, e.g., as a traditional table, on maps, time lines, and using pie and other charts. Detailed information on individual result entities can be seen on their *instance pages*.

Fig. 1 illustrates an example use case of the portal where the user wants to find novels that match specific criteria. The user first chooses the novels perspective. The user then makes selections on three facets: genre=*romance novels*, setting=*castles* and characters=*nobility*. The and results and facet hit counts are automatically updated after the makes a choice. The user can finally choose a novel that looks appealing and open its instance page to see aggregated information regarding the novel.

Fig. 2 illustrates another use case. Here the user has chosen the publications perspective willing to analyze the evolution of novel themes in time. The user has made facet selection publications=*original*, i.e., not translations, type=*novels*, and language=*Finnish*. The visualization tab of *Annual themes and keyword consists* of two parts: The upper part shows the number of publications per the top 10 themes throughout the years. The lower visualization has the same idea

---

[6] Available at: http://kirjasampo.fi

[7] https://kirjastot.fi

[8] Statistics (in Finnish) available at: https://www.kirjasampo.fi/fi/kirjasammon-tilastot-2021
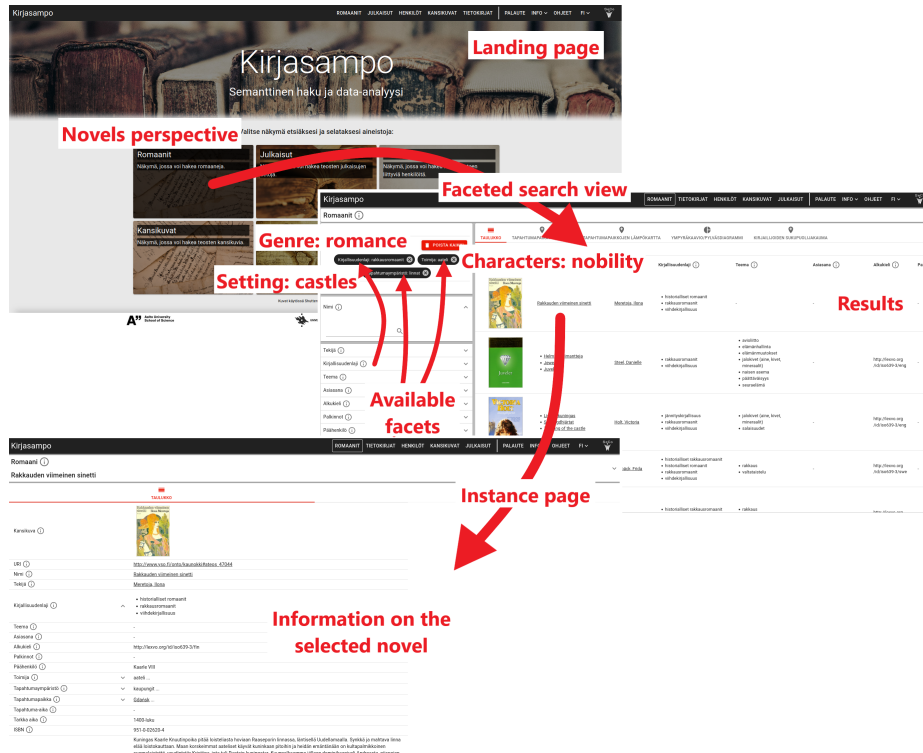
Fig. 1. An example use case of the portal: Finding a novel fulfilling criteria.

but for keywords, which are used in the BookSampo KG to supplement themes when no appropriate theme entity is found for a novel.

Using BookSampo is demonstrated in more detail in this online video[9].

4 Discussion and Future Work

The presented UI utilizes the possibilities of semantic web technologies to enhance the search, browsing, and data analysis experience for the end-user. The end-goal is to publish this new UI alongside the original portal. This requires some cleaning of the underlying KG, which is already underway. Making the portal available for public use opens up the possibility of gathering feedback from end-users on how the portal and its functionalities can be improved or changed to better accommodate the end-users' needs. We also plan to publishing the KG openly for data analytic research.

The underlying KG could also be enriched with data from other KGs in the Cultural Heritage (CH) domain, such as the BiographySampo KG [4], which
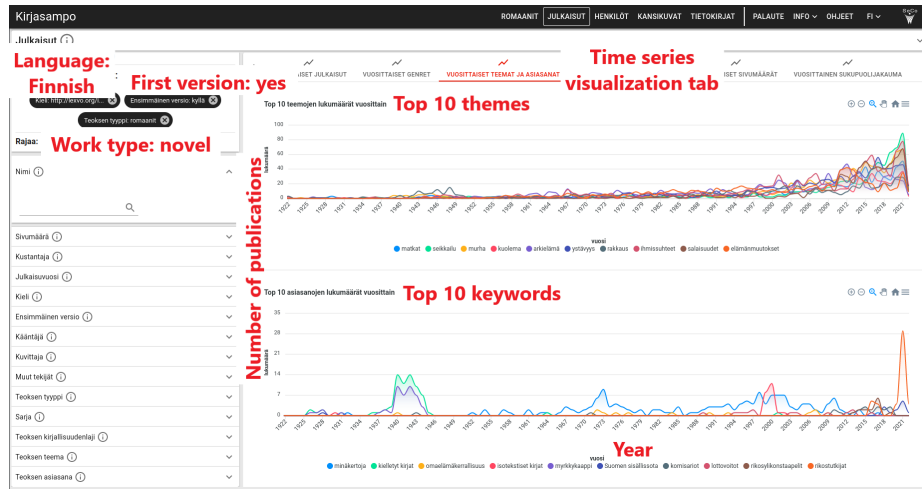

Fig. 2. A time series visualization of the top 10 themes and keywords.

already contains over 700 links to people (e.g., authors) in the BOOKSAMPO KG. This would allow the users to better explore the world of Finnish literature and related people by providing linking external of the BOOKSAMPO KG.

References

1. Haslhofer, B., Isaac, A., Simon, R.: Knowledge graphs in the libraries and digital humanities domain. arXiv preprint arXiv:1803.03198 (2018)
2. Hyvönen, E., Viljanen, K., Tuominen, J., Seppälä, K.: Building a National Semantic Web Ontology and Ontology Service Infrastructure – The FinnONTO Approach. In: Proceedings of the ESWC 2008, Tenerife, Spain. pp. 95–109. Springer (2008)
3. Hyvönen, E.: Digital Humanities on the Semantic Web: Sampo Model and Portal Series. Semantic Web – Interoperability, Usability, Applicability (2022), http://semantic-web-journal.org/content/digital-humanities-semantic-web-sampo-model-and-portal-series, accepted
4. Hyvönen, E., Leskinen, P., Tamper, M., Rantala, H., Ikkala, E., Tuominen, J., Keravuori, K.: BiographySampo – Publishing and enriching biographies on the Semantic Web for digital humanities research. In: Proceedings of the 16th Extended Semantic Web Conference (ESWC 2019). pp. 574–589. Springer (2019)
5. Ikkala, E., Hyvönen, E., Rantala, H., Koho, M.: Sampo-UI: A Full Stack JavaScript Framework for Developing Semantic Portal User Interfaces. Semantic Web – Interoperability, Usability, Applicability **13**(1), 69–84 (2022). https://doi.org/10.3233/SW-210428
6. Mäkelä, E., Hypén, K., Hyvönen, E.: BookSampo—lessons learned in creating a semantic portal for fiction literature. In: Proc. of ISWC-2011, Bonn, Germany. Springer (2011)
7. Mäkelä, E., Hypén, K., Hyvönen, E.: Fiction Literature as Linked Open Data – the BookSampo Dataset (2013). https://doi.org/10.3233/SW-120093