Visualizing Literary Linked Data for Public Library Users in the New User Interface for BookSampo – Finnish Fiction Literature on the Semantic Web

Annastiina Ahola[1], Eero Hyvönen[1,2]

[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

The BookSampo Linked Data (LD) portal was deployed in 2011 by the Public Libraries of Finland and has today nearly 1.6 million annual users. Its large knowledge graph (KG) covers virtually all Finnish Fiction literature but has not been fully exploited in Digital Humanities. This paper discusses how the KG can be used for literary search, data exploration, and research by presenting a new BookSampo user interface (UI) based on faceted semantic search and browsing with seamlessly integrated data-analytic visualization tools. This application makes it possible for the first time to analyze the BookSampo data in versatile ways without programming skills. The analysis results presented suggest interesting spatial, temporal, and topical trends in how the Finnish fiction literature has evolved during the last decades.

1. Introduction

BookSampo[1] [1, 2] provides information on virtually all fiction literature published in Finland since mid 19th century. Its contents are based on rich semantic descriptions of books and theirs contexts using Linked Data (LD) that originates from multiple heterogeneous data sources. BookSampo is an application instance of the more general "Sampo Model"[2] for LD publishing and series of semantic portals in use[3] in Finland and beyond [3].

BookSampo is used by library users and librarians for finding literary works of interest and related contextual information. The original Drupal-based UI in use since 2011 provides traditional text search engines for finding records and then related contents as links for data exploration [4]. However, the full potential of the underlying KG or searching, exploring, and for data analytic research has not been exploited.

To facilitate this, this paper presents a new semantic user interface (UI) for for the BookSampo knowledge graph (KG). First in Section 2 the Sampo model underlying our work is overviewed.

0009-0008-6369-4712 (A. Ahola); 0000-0003-1695-5840 (E. Hyvönen)
© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

[1]Portal: http://kirjasampo.fi; research homepage: https://seco.cs.aalto.fi/applications/kirjasampo/
[2]The model is called "Sampo" according to the Finnish epic Kalevala, where Sampo is a mythical machine giving riches and fortune to its holder, a kind of ancient metaphor of technology according to the most common interpretation of the concept.
[3]See https://seco.cs.aalto.fi/applications/sampo/ for Sampos, links, videos, and publications.

CEUR-WS.org/Vol-3508/paper1.pdf

After this the BookSampo KG is presented (Section 3). Based on using the Sampo-UI framework [5, 6], section 4 explains the new semantic portal with examples of using the system. In conclusion, related works are over-viewed, contributions of the new UI are discussed, and next steps ahead are outlined. This paper extends substantially our earlier short papers of the new project [7] and on using the BookSampo [8].

2. Sampo Model: Publishing and Studying Linked Data

Table 1
Sampo Model Principles P1–P6

P1	Support collaborative data creation and publishing
P2	Use a shared open ontology infrastructure
P3	Make clear distinction between the LOD service and the user interface (UI)
P4	Provide multiple perspectives to the same data
P5	Standardize portal usage by a simple filter-analyze two-step cycle
P6	Support data analysis and knowledge discovery in addition to data exploration

The Sampo model[4] [3] is a consolidated set of principles listed in Table 1 for collaborative LOD publishing and creating portals. The model is based on the Semantic Web standards[5] [9] and best practices of the W3C for Linked Data publishing [10, 11].

Principles P1–P3 lay a foundation for developing LOD services. The model is based on the idea of collaborative content creation (P1)[6]. The data is aggregated from local data silos into a global service, based on a shared ontology and publishing infrastructure (P2). The model supports the idea of separating the underlying Linked Data service *completely* from the user interface via a SPARQL API (P3). This arguably simplifies the portal architecture and the data service can be opened for data analysis research in Digital Humanities.

The idea of principles P4–P6 is to "standardize" the UI logic of Sampo portals to be created on top of a LD service SPARQL endpoint. The goal is to make the portals easier to use and implement. Principle P4 articulates the idea of providing different thematic *application perspectives* by re-using the data service. They are used by a two-step cycle for research (P5): First the focus of interest, the target group, is filtered out using faceted semantic search [12, 13, 14]. Second, the target group is visualized or analyzed by using ready-to-use data analytic tools of the application perspectives. A novelty of the model is to support data analysis, visualization, and knowledge discovery with seamlessly integrated tooling, even solving research problems using AI [15] (P6).

To create Sampo portals, the Sampo-UI framework [5] has been designed and implemented. It has been used for developing all ca. 15 Sampo systems since 2018, including the new BookSampo portal, suggesting practical feasibility of the Sampo model [6].

[4]The name "Sampo" comes from the Finnish epic Kalevala, where Sampo is a mythical machine giving riches and fortune to its holder, a kind of ancient metaphor of technology.
[5]https://www.w3.org/standards/semanticweb/
[6]In our case the collaborators are institutions rather than individual people.

3. BookSampo Knowledge Graph

The user interface (UI) on top of the underlying knowledge graph (KG) in a SPARQL endpoint has been created using traditional search and data exploration methods, and the full potential of the KG—nearly 9 million triples today—has not been fully utilized: the data covers all Finnish fiction literature and beyond and is interesting from a Digital Humanities (DH) research perspective, too. Table 2 lists the number of instances of different entity types in the data from back in 2013 and today.

Class (Type)	Instances
Literary works	93,000
Editions	127,000
Book covers	27,000
Fictional characters	19,000
Contemporary reviews	15,000
Literary series	2,900
Literary awards	2,700
Literary award series	200
People (e.g. authors)	29,000
Author's pictures	2,600
Publishers	2,600

Class (Type)	Instances
Literary works	215,000
Editions	222,000
Book covers	119,000
Fictional characters	49,000
Contemporary reviews	15,000
Literary series	8,900
Literary awards	6,400
Literary award series	300
People (e.g. authors)	64,000
Author's pictures	4,200
Publishers	5,500

Table 2
Instance counts in 2013 [1] (left) vs. 2023 (right)

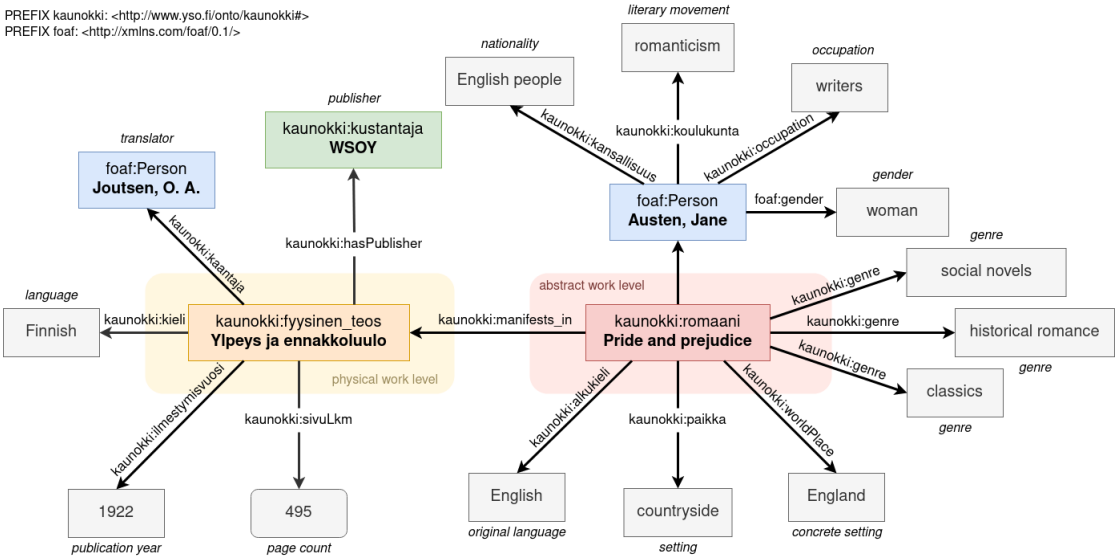

Figure 1: An example of how a novel is modeled in the BookSampo KG.

Fig. 1 illustrates how the novel *Pride and prejudice* is modeled in the BookSampo KG. The

`kaunokki:romaani` entity represents the abstract work level of the novel. That entity has links to the entity of the author *Austen, Jane* and the `kaunokki:fyysinen_teos` entity representing the Finnish edition of the work translated by *Joutsen, O. A.*. This Finnish edition further has the link to the publisher entity of the *WSOY* publishing house.

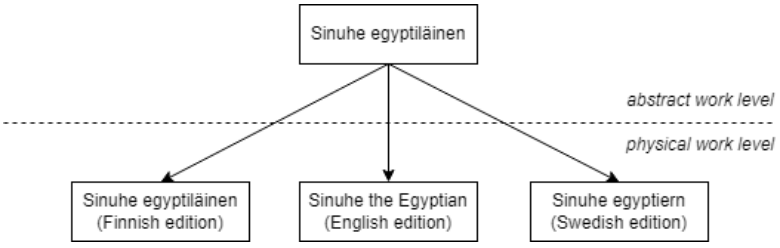

Figure 2: The abstract and physical work levels for Mika Waltari's novel *Sinuhe Egyptiläinen*.

The BookSampo data divides works into two levels: abstract and physical work levels. The data model is based on the FRBRoo model [16] but simplified to the aforementioned two levels. The abstract work level is equivalent to the *work* level in the FRBRoo model while the physical work level represents what would be the *manifestation* level in the FRBRoo model. In practice the abstract work level deals with information that is shared between all editions of a work. The physical work level on the other hand contains edition-specific information, e.g., number of pages, and publisher, that are specific to the edition. In the case of the BookSampo data this edition-specific information is often recorded for the first editions in all relevant languages. That is to say, an average work will have a physical work level entities for at least its first edition in the original language as well as for the possible first editions of translations of the work into Finnish and/or Swedish. If a work is translated again later in time, an additional physical work level entity might be added for the new updated translation. Fig. 2 illustrates the split for Mika Waltari's novel *Sinuhe Egyptiläinen*.

4. New BookSampo User Interface

This section first present the standard model of Sampo-UI for designing and implementing semantic portals. After this, it is shown how the model was used in the BookSampo portal.

Generic Sampo-UI model The new BookSampo User Interface (UI) is built using the Sampo-UI framework [5] based on the Sampo model [3] introduced in Section 2. A Sampo-UI interface contains a landing page with a set of "application perspectives" to choose from as depicted in Fig. 3. Each perspective provides the end-user with a faceted semantic search view to filter out individuals of a class (e.g., Novel, Person, Place, etc.) related to the perspective. After finding a set on instances (individuals) of interest, the user can either study and analyze them 1) one by one, 2) analyze sets of individuals together (e.g., statistically, on maps, or on timelines), or 3) explore the data by browsing based on internal links within the portal or links to external data sources such as Wikipedia or or other Sampo portals. Data-analytic view and visualization can be customized for each application perspective for groups of instances and individual instances

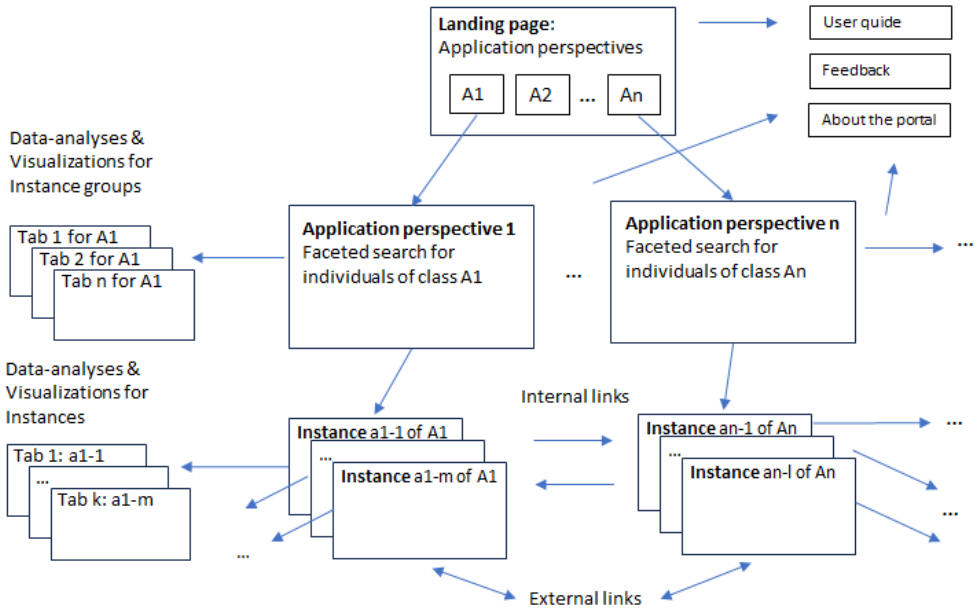

Figure 3: Navigational page structure of a portal based on Sampo-UI.

separately. From all pages links to a user guide, feedback channel to developers, and a page about the portal are provided (cf. the upper right corner in the figure).

A Sampo portal is implemented by configuring the pages types depicted in Fig. 3, i.e., the landing page, application perspective pages, and the instance pages. This is done using JSON (JavaScript Object Notation) configuration files [6]. The Sampo-UI framework offers ready-to-use components to be used in portal pages that can be added through the configuration files without the need for heavy coding. The components can easily be expanded upon by adding new mapping functions and expanding upon configuration options passed to the components.

Structure of the UI The user of BookSampo first lands on the **landing page** (shown in Fig. 4) when opening the portal. The portal consists of five *application perspectives* based on corresponding classes:

1. *Novels.* This perspective deals with the abstract work level of novels.
2. *Publications.* This perspective deals with the physical work level of *all* works.
3. *People.* This perspective deals with authors and other people related to literature, e.g., illustrators, translators, and reviewers.
4. *Covers.* This perspective deals with book covers and information related to them.
5. *Nonfiction books.* This perspective deals with the abstract work level of nonfiction books.

All of the perspectives query the data from the same SPARQL endpoint of BookSampo KG. Selecting a perspective represents the data from that perspective. For example, choosing the

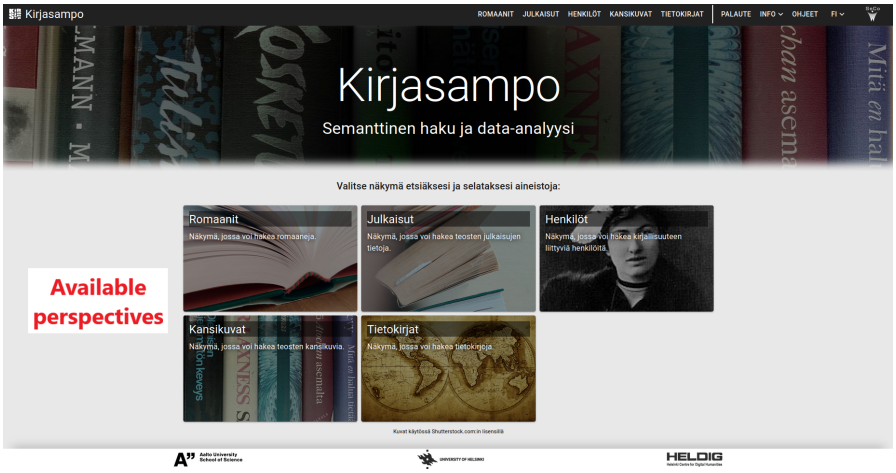

Figure 4: The landing page of the BookSampo UI with five application perspectives.

Novels perspective shows the data through novels. The search view of that perspective lists all novel entities in a list with links to the rest of the KG as their properties.

The basis for choosing the aforementioned perspectives was to cover all aspects of Finnish literature as comprehensively as possible with few perspectives. *Novels* perspective was chosen as the perspective to cover fictional books due to novels being the largest subgroup of fictional literary works in the data. *Nonfiction books* was chosen to supplement the *Novels* perspective with the nonfictional works, although in the data it only represents a non-comprehensive subset of nonfiction published in Finland. The *Publications* perspective covers all literary works on the physical work level. The split between 1) novels and nonfiction books and 2) publications follows the split made in the original BookSampo data [1] as introduced in Section 3.

To supplement the data on literary works, the *People* perspective was added to provide information on all people relevant to the presented data on literary works, whether it be the authors behind the books or other people relevant to them like the illustrators and translators of works. As the BookSampo KG also includes data on contemporary reviews, information of reviewers is also included in this perspective. To finish off the available perspectives, the *Covers* perspective was added due to the popularity of the book cover search function on the original BookSampo Portal, the search capabilities of which could be even further improved with the use of faceted search.

Clicking on a card for an application perspectives leads to that particular perspective's **faceted search view** (see Fig. 5). The faceted search view consists of three key elements: 1) the facet menu on the left, 2) the results view on the right, and 3) the different visualization tabs on top of the results view.

The facet menu includes all the available facets that can be used to filter the data in that particular perspective. The Sampo-UI framework offers various types of facets for different types of data. The BookSampo Portal utilizes three different facets types:

1. *Checkbox facet.* A facet for filtering results by selecting one or multiple checkboxes for

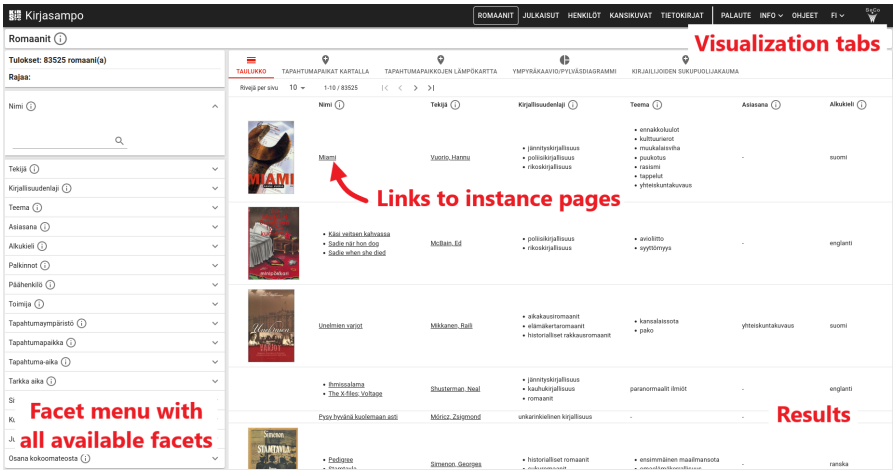

Figure 5: The faceted search view of the *Novels* perspective in the BookSampo UI.

wanted property value entities. Results are automatically updated when a checkbox is checked. Selecting multiple checkboxes works in a disjunctive way: Selecting both the genres *romance* and *thriller* for the genre facet would return works that belong to either (or both) genre. If the property values have a hierarchical structure, e.g., yearly literary awards (*Finlandia Prize 2022*) and the award series they belong to (*Finlandia Prizes*), the facet can be configured to show the entities hierarchically as well.

2. *Integer range facet.* A facet for filtering results by limiting the integer range a property's value should be in, e.g. searching for works that have a *page count* in the range of 200–300. The facet is applied by pressing the 'apply' button.

3. *Text facet.* A facet for searching for results based on text string, e.g. searching for works by their *names*. The facet is applied after the user presses enter.

The results are shown in a table format on the right side of the screen. Each of the rows represents one entity. The different columns represent the different properties and property values these entities have. Column values that are underlined denote links to more information about that particular entity. By default the result set includes all entities that match the type of the application perspective with no filters applied. The results are automatically updated when any facets have been applied.

The links in the results table lead to the **instance pages** (shown in Fig. 6) of entities. Instance pages aggregate all the information about that particular object in the same page. This includes the information shown about the entity in the table view as well as possible further information not deemed relevant to be included in the table view. In the BookSampo Portal the general choice was to include information that could be used as facets as columns in the table view and leave the rest of the relevant information to the instance pages.

Similarly to the faceted search view, the instance page view can have multiple tabs for different ways of visualizing the data. These tabs depend on the type of the entity in question. Novels

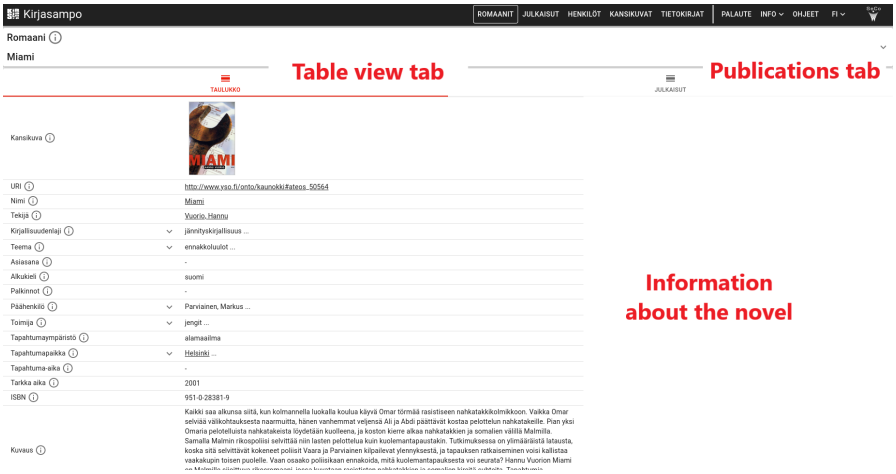

Figure 6: The instance page of Vuorio Hannu's novel *Miami* in the BookSampo UI.

and nonfiction books, for example, have a specific tab for showing detailed information about the different publications of that particular work (shown in Fig. 7) that exist in the data.

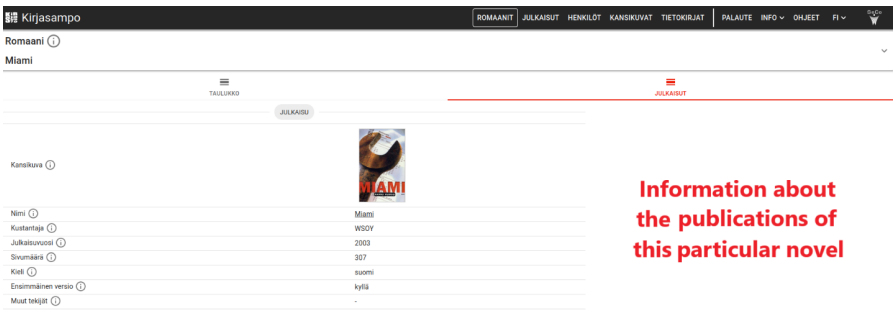

Figure 7: The editions list of Hannu Vuorio's novel *Miami*; the singular edition of the work is shown.

Visualizing the data with integrated data-analytic tools The various tabs available in the faceted search views and instance pages of application perspectives offer the user different ways of looking at the data as well as analyzing the data. The different visualization types available can be roughly split into three different categories:

1. *Pie/bar charts.* Charts that show the ratio of property values in comparison to each other, e.g., the top genres of novels (shown in Fig. 8). The component behind these visualizations

is created with the ApexCharts[7] library.

2. *Maps.* Charts that show entities on a map based on some location information related to the entity, e.g., settings of novels on a map (shown in Fig. 9). The components behind map visualizations are created with the Leaflet[8] and deck.gl[9] libraries.

3. *Time series.* Charts that show the evolution of entities or some of their properties as a function of time, e.g., the evolution of average page counts throughout years (shown in Fig. 10). The components behind these visualizations are created with the ApexCharts and AmCharts[10] libraries.

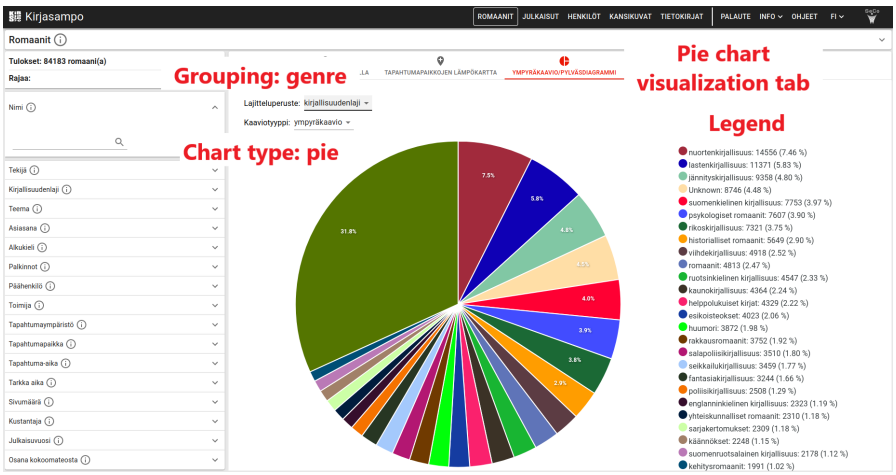

Figure 8: The top genres of novels visualized as a pie chart.

The Sampo-UI framework offers some ready-to-use visualization types for all of these categories. To better utilize the potential of the data, some custom visualizations were developed. These visualizations use the libraries already included in the framework, but have custom configuration options, mappers and/or data processing functions to expand on the possible visualization types available in the base Sampo-UI framework. The inspiration behind these components came from the data itself, where no existing visualization type could readily be used to visualize an aspect that could potentially turn out to be interesting.

Two examples of such components are shown in Fig. 11 and Fig. 12. Fig. 11 shows a deck.gl-based component for visualizing the gender ratio of authors, who have written novels with a location as a novel's setting. The circles on the map represent different locations, where the size of the circle is determined by the number of novels with that location as their setting and the color of the circle represents the gender ratio for that location. Red circles indicate a

[7] https://apexcharts.com/
[8] https://leafletjs.com/
[9] https://deck.gl/
[10] https://www.amcharts.com/

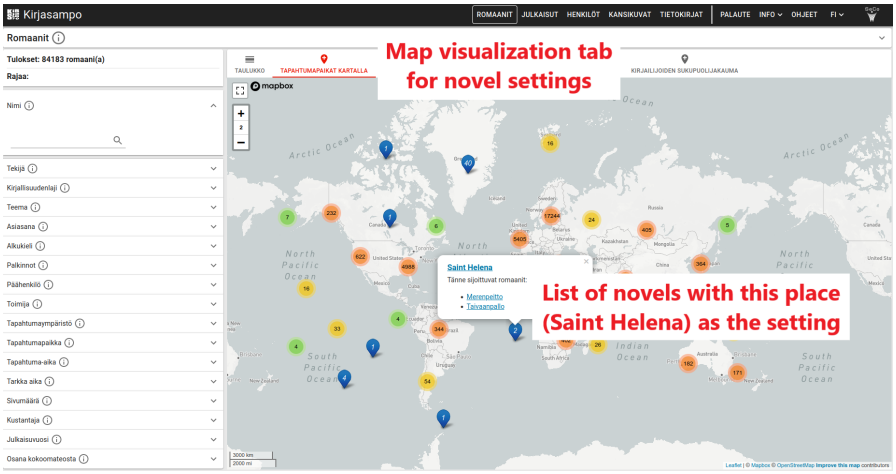

Figure 9: The settings of novels on a map.

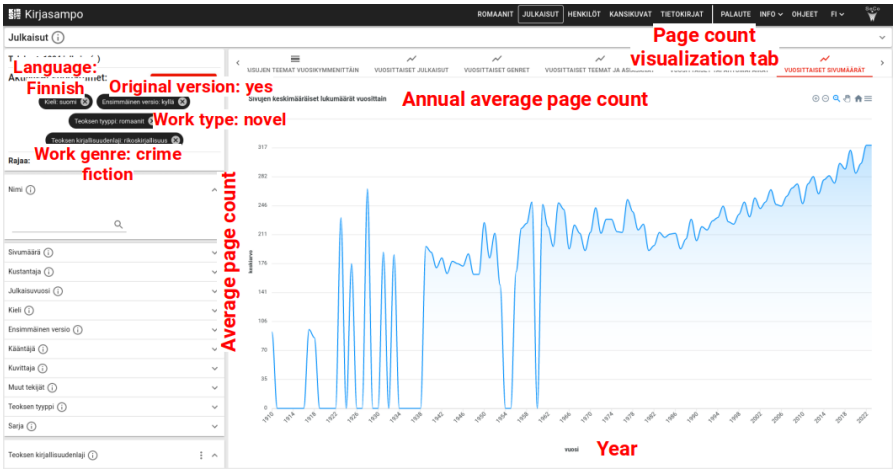

Figure 10: The evolution of average page count of publications for Finnish crime fiction novels.

strong female author majority and blue circles a strong male author majority, with the purplish shades in between indicating a more even split between female and male authors. In Fig. 11, for example, interestingly the major locations in the UK have at least a slight female author majority, while Germany on the other hand has more locations with a male author majority. With the faceted search tools the UI offers, the user could easily explore in more detail whether, for example, the top genres for these locations could play a part in the difference.

Fig. 12 is based on an existing ApexCharts time series component configuration, but with new data preprocessing functions to enable it to handle dynamic sets of series. This enables the visualization component to render the top themes and keywords and their evolution throughout the years for publications dynamically based on the applied facets instead of having to

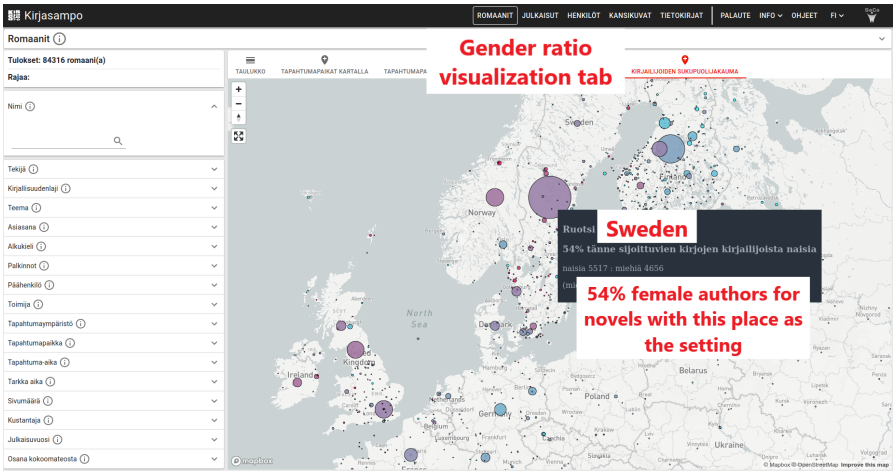

Figure 11: A map showing the gender ratio of the authors whose novels have a place as their setting.

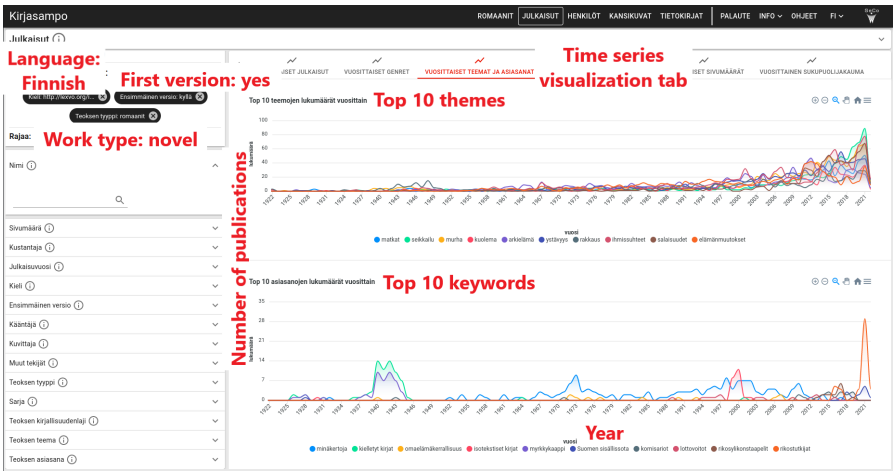

Figure 12: A timeline visualization showing the top themes and keywords of original Finnish novels.

predetermine the visualized themes and keywords. In the bottom visualization for the top 10 keywords, for example, there is an interesting visible spike for two keywords – *kielletyt kirjat*[11] and *myrkkykaappi*[12] – during the late 1930s to 1940s, when the Winter and Continuation Wars were fought in Finland.

The general idea behind the inclusion of visualizations is that they should visualize the entities being filtered in that particular perspective. For example, in the *Novels* perspective all the visualizations mainly deal with the novels and their properties. Ideally the set of the

[11] Finnish for 'banned books'.

[12] Lit. 'poison cabinet'. This is a term used for a cabinet in certain libraries that would be used to store controversial/banned books.

visualized entities should be the exact same result set as shown in the table view of the results. The problem, however, arises with properties having varying annotation coverage. This is especially apparent with the map visualizations, where the ratio of the set of visualized entities to the whole result set can become low, e.g., only around 12,000 people out of 62,000 people having an annotated place of birth. The problem is also further exacerbated by issues with the geographical location data in the BookSampo KG, where a single location can have multiple different latitudes and longitudes listed in the data. In these cases those locations are just filtered out of the results to prevent them from skewing the visualization by showing the 'same location' and thus instance counts at multiple different locations.

The choice on whether to even include a visualization in the first place was based on how potentially interesting the data itself could be: The settings of novels as well as the places of birth and death could potentially show some interesting patterns and correlations between factors, but on the other hand something like the place of education like won't offer much due to the extremely low annotation coverage (less than 2,000 people out of 62,000 people) as well as the most common annotations largely being Finnish cities with major universities.

For pie/bar charts this problem with coverage is somewhat easier to mitigate by being able to visually show the ratio of known and unknown values by including a slice with the label 'Unknown'. This also helps the user more easily assess how well the annotated values might represent the whole data set. For map visualizations and most of the time series visualizations this ratio of known to unknown isn't as clear unless the user manually counts together the instance counts in the visualizations.

5. Discussion

Related Works Linked Data and ontologies have been used in libraries [17], museums [18, 19, 20], and archives [21, 22]. Using LD is advocated by major library organizations, such as IFLA[13] and OCLC [14], and several libraries provide their collections as data in this form [23]. LD has been used in building infrastructures, such as ARIADNEplus [15] for archaeology, Linked Art[16] in the U.S., and in local efforts in Italy [24], the U.K. [25], Spain [26], and Finland [27] to list a few examples. Cultural Heritage and DH have become a major application domain for LD technologies [28, 29]. The focus in related research on portal UIs has been usually on (explorative) search and browsing [4]. In contrast, Sampo systems have a strong focus on data analysis and visualizations integrated seamlessly with (faceted) search and browsing.

Evaluation The new BookSampo UI has been available mostly to members of the Semantic Computing Research Group as well as people from the Finnish Public Libraries while the most important issues of the underlying KG are being corrected. The UI has been improved based on the wants and needs of the current users, e.g. by including additional facets the users felt the UI was missing and by implementing new custom visualizations to better visualize aspects of the data, but hasn't yet been evaluated with outside testers. Based on previous in-use semantic

[13] https://www.ifla.org/references/best-practice-for-national-bibliographic-agencies-in-a-digital-age/service-delivery/linked-open-data/
[14] https://www.oclc.org/research/areas/data-science/linkeddata/linked-data-overview.html
[15] https://ariadne-infrastructure.eu/
[16] https://linked.art/

portals utilizing the Sampo-UI framework, the framework is however suggested to have good usability and scalability for the end user [30, 5, 31].

In terms of the evaluation of the contents of the BookSampo Portal itself, the Sampo-UI framework's structure enables easily developing future improvements. The Sampo-UI framework offers various ways to extend and customize the portal through things like custom components based on the needs of the users, which was already done during development based on the comments from the members of the research group as well as the people from the public libraries. Therefore possible problems or lacking features could be implemented and/or fixed in the future when the portal is made public.

The BookSampo KG has already shown itself to be interesting from the perspective of DH research [32], which the new UI could potentially make more accessible to both non-academic people as well as researchers not previously familiar with the data and its potential. The new UI offers ways to explore and browse the data without the need for much technical knowledge or skills in query or programming languages like SPARQL and Python that are often used when conducting research with these kinds of data sets.

The ease of use, however, also means less ability for the user to affect the data being visualized. For example, the data on geographical locations has its problems with the excess longitude and latitude values as well as a missing hierarchy system for connecting cities and towns to the countries they are in. Issues like these would greatly benefit from the ability to preprocess the data before visualizing it, which would be available to the user if they were working directly with SPARQL and Python instead.

Another problem arises from the nature of simply how things are annotated. The annotations are done by humans who naturally might have their own ways and preferences of how they annotate things. For example, one person might annotate a book's setting as 'Helsinki' while another might annotate it as 'Finland'. Both of these choices could be correct on a level, but the final choice is dependant on the annotator. With geographical locations, adding hierarchy and visualizing things on a country level could lead to most accurate visualizations in some cases, but with something like themes and genres, aggregating things to an upper shared concept might not be as easy as stories with similar elements can still be described in different ways.

Future Work The new UI offers the users a new and more intuitive way to explore the BookSampo KG compared to the old portal with its limited search functions. With the integrated data-analytic tools the users can also easily analyze the data they are presented without having to learn languages like Python to work with the result data from SPARQL queries. The new UI could potentially serve as an intermediary step in DH research. The underlying BookSampo KG has a lot of potential to be used in literary DH research, but has not widely been used for it yet. With the new UI researchers could easily search, explore, and analyze the underlying data without needing to be familiar with technology like SPARQL or Python that would be needed otherwise. In addition to not needing to be familiar with the technology, the UI could be helpful in just narrowing down the are of interest to be researched. In comparison to trying to scour through the KG itself with its nearly 9 million triples, finding interesting topics or questions is be easier with the new UI, where one can expand on search in an iterative way by adding new restrictions on the results through the facets as you go.

Due to the nature of the Sampo-UI framework, the portal could easily be extended in the future. The BookSampo KG includes data on other types of works, e.g., poems and short story

collections, that could easily be added to the portal as perspectives after assessing the quality of that data and the annotations related to them. The components and visualizations present in the portal could be expanded as well based on the needs and wishes of users and their feedback. The data could also be enriched by linking it to other linked data sources, e.g., BiographySampo KG [33], which includes data on over 700 people present in the BookSampo KG.

The new portal presented in this paper is planned to be opened for public use in autumn 2023. Before this can be done, the data issues need are being addressed by the Public Libraries of Finland. We also plan to publish (most of) the data for open use with the CC BY 4.0 license at the Linked Data Finland platform[17].

Acknowledgements Thanks to Matti Sarmela, Kaisa Hypén, and Tuomas Aitonurmi for their collaborations as well as providing a newer version of the BookSampo KG to be used in the development of the new user interface. This project was funded by the Aalto University; Computing resources of the CSC – IT Center for Science were used.

References

[1] E. Mäkelä, K. Hypén, E. Hyvönen, Fiction Literature as Linked Open Data – the BookSampo Dataset, 2013. doi:10.3233/SW-120093.

[2] E. Mäkelä, K. Hypén, E. Hyvönen, BookSampo—lessons learned in creating a semantic portal for fiction literature, in: Proc. of ISWC-2011, Bonn, Germany, Springer, 2011.

[3] E. Hyvönen, Digital Humanities on the Semantic Web: Sampo Model and Portal Series, Semantic Web – Interoperability, Usability, Applicability 14 (2023) 729–744. doi:10.3233/SW-223034.

[4] G. Marchionini, Exploratory search: from finding to understanding, Communications of the ACM 49 (2006) 41–46. doi:10.1145/1121949.1121979.

[5] E. Ikkala, E. Hyvönen, H. Rantala, M. Koho, Sampo-UI: A Full Stack JavaScript Framework for Developing Semantic Portal User Interfaces, Semantic Web – Interoperability, Usability, Applicability 13 (2022) 69–84. doi:10.3233/SW-210428.

[6] H. Rantala, A. Ahola, E. Ikkala, E. Hyvönen, How to create easily a data analytic semantic portal on top of a SPARQL endpoint: introducing the configurable Sampo-UI framework, 2013. Paper submitted for VOILA 2023 workshop for evaluation.

[7] E. Hyvönen, A. Ahola, E. Ikkala, BookSampo fiction literature knowledge graph revisited: Building a faceted search interface with seamlessly integrated data-analytic tools, in: Theory and Practice of Digital Libraries (TDPL 2022), Accelerating Innovations Track, Springer, 2022.

[8] A. Ahola, E. Hyvönen, H. Rantala, A user interface model for digital humanities research: Case BookSampo – Finnish Fiction Literature on the Semantic Web, in: Proceedings of ESWC 2023, poster and demo papers, Springer, 2023.

[9] P. Hitzler, M. Krötzsch, S. Rudolph, Foundations of Semantic Web technologies, CRC press, 2010.

[10] T. Heath, C. Bizer, Linked Data: Evolving the Web into a Global Data Space (1st edition),

[17] https://ldf.fi

Morgan & Claypool, Palo Alto, California, 2011. URL: http://linkeddatabook.com/editions/1.0/.
[11] E. Hyvönen, Publishing and using cultural heritage linked data on the Semantic Web, Morgan & Claypool, Palo Alto, California, 2012.
[12] E. Hyvönen, S. Saarela, K. Viljanen, Application of ontology-based techniques to view-based semantic search and browsing, in: Proceedings of the First European Semantic Web Symposium, Springer, 2004.
[13] D. Tunkelang, Faceted search, Morgan & Claypool Publishers, CA, USA, 2009.
[14] Y. Tzitzikas, N. Manolis, P. Papadakos, Faceted exploration of RDF/S datasets: a survey, Journal of Intelligent Information Systems 48 (2017) 329–364. doi:10.1007/s10844-016-0413-8.
[15] E. Hyvönen, Using the Semantic Web in Digital Humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery, Semantic Web – Interoperability, Usability, Applicability 11 (2020) 187–193.
[16] P. Riva, M. Doerr, M. Zumer, FRBRoo: enabling a common view of information from memory institutions, in: World Library and Information Congress: 74th IFLA General Confrence and Council, 2008.
[17] B. Haslhofer, A. Isaac, R. Simon, Knowledge graphs in the libraries and digital humanities domain, arXiv preprint arXiv:1803.03198 (2018).
[18] L. Aroyo, N. Stash, Y. Wang, P. Gorgels, L. Rutledge, CHIP demonstrator: Semantics-driven recommendations and museum tour generation, in: The Semantic Web, Springer, 2007, pp. 879–886. doi:10.1007/978-3-540-76298-0_64.
[19] G. Schreiber, A. Amin, L. Aroyo, M. van Assem, V. de Boer, L. Hardman, M. Hildebrand, B. Omelayenko, J. van Osenbruggen, A. Tordai, J. Wielemaker, B. Wielinga, Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator, Journal of Web Semantics 6 (2008) 243–249. doi:10.1016/j.websem.2008.08.001.
[20] P. Szekely, C. A. Knoblock, F. Yang, E. E. Fink, S. Gupta, R. Allen, G. Goodlander, Publishing the data of the Smithsonian American Art Museum to the linked data cloud, International Journal of Humanities and Arts Computing 8 (2014) 152–166. URL: http://usc-isi-i2.github.io/papers/szekely14-ijhac.pdf.
[21] S. Van Hooland, R. Verborgh, Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata, Facet Publishing, 2014. doi:10.1080/00048623.2016.1162277.
[22] M. Hallo, S. Luján-Mora, A. Maté, J. Trujillo, Current state of linked data in digital libraries, Journal of Information Science 42 (2016) 117–127. doi:10.1177/0165551515594729.
[23] E. T. Mitchell, Library linked data: early activity and development, ALA TechSource Chicago, IL, 2016.
[24] V. A. Carriero, A. Gangemi, M. L. Mancinelli, L. Marinucci, A. G. Nuzzolese, V. Presutti, C. Veninata, ArCo: The Italian cultural heritage knowledge graph, in: The Semantic Web – ISWC 2019, Springer, 2019, pp. 36–52. doi:10.1007/978-3-030-30796-7_3.
[25] Y. Lei, V. Lopez, E. Motta, V. Uren, An infrastructure for semantic web portals, Journal of Web Engineering 6 (2007) 283–308. URL: https://journals.riverpublishers.com/index.php/JWE/article/view/4105.
[26] F. Hernández, L. Rodrigo, J. Contreras, F. Carbone, Building a cultural heritage ontology

for Cantabria, in: Annual Conference of CIDOC, 2008, pp. 1–14. URL: https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2018/12/64_papers.pdf.

[27] E. Hyvönen, K. Viljanen, J. Tuominen, K. Seppälä, Building a National Semantic Web Ontology and Ontology Service Infrastructure – The FinnONTO Approach, in: Proceedings of the ESWC 2008, Tenerife, Spain, Springer, 2008, pp. 95–109.

[28] A. Bikakis, E. Hyvönen, S. Jean, B. Markhoff, A. Mosca, Special issue on semantic web for cultural heritage, Semantic Web – Interoperability, Usability, Applicability 12 (2021). doi:10.3233/SW-210425.

[29] M. Zeng, C. Sula, K. Gracy, E. Hyvönen, V. M. A. Lima, JASIST special issue on digital humanities (DH), Journal of the Association for Information Science and Technology (JASIST) (2021) 1–5. doi:10.1002/asi.24584.

[30] T. Burrows, N. B. Pinto, M. Cazals, A. Gaudin, H. Wijsman, Evaluating a semantic portal for the "Mapping Manuscript Migrations" project, DigItalia 15 (2020) 178–185.

[31] J. English, M. Hearst, R. Sinha, K. Swearingen, K. Lee, Flexible search and navigation using faceted metadata, Technical Report, Technical report, University of Berkeley, School of Information Management and Systems, 2002.

[32] T. Peura, Suomeksi yli rajojen. Kvantitatiivinen tutkimus suomenkielisten romaanien monimuotoisuudesta 1970-2020, Master's thesis, University of Helsinki, Department of Digital Humanities, Helsinki Centre for Digital Humanities (HELDIG), 2023.

[33] E. Hyvönen, P. Leskinen, M. Tamper, H. Rantala, E. Ikkala, J. Tuominen, K. Keravuori, BiographySampo – Publishing and enriching biographies on the Semantic Web for digital humanities research, in: Proceedings of the 16th Extended Semantic Web Conference (ESWC 2019), Springer, 2019, pp. 574–589.