

Minutes of Multilateralism on the Semantic Web – League of Nations Sampo (1920–1946) Portal for Digital Humanities Research

Eero Hyvönen^{2,1}, Petri Leskinen^{1,2}, Grégoire Mallard¹, Pierre Etienne Bourneuf¹, Alexandre Lionnet^{3,1}, Blandine Blukacz-Louisfert⁴, Davide Rodogno¹, and Florian Cafiero^{3,1}

¹ Geneva Graduate Institute, Center for Digital Humanities and Multilateralism, Geneva, Switzerland

² Aalto University, Department of Computer Science, Semantic Computing Research Group (SeCo), Finland

³ École nationale des chartes, Paris, France

⁴ UN Library & Archives Geneva, Geneva, Switzerland

Abstract. This paper demonstrates, by presenting a new Linked Open Data (LOD) service and semantic portal LEAGUE OF NATIONS SAMPO, how semantic web technologies can be used in Digital Humanities (DH) research to search, browse, and analyze assembly minutes of multilateral international organizations and their representatives.

Keywords: knowledge graphs · digital humanities · data analysis

1 Introduction

International organizations, such as the League of Nations, United Nations, and Inter-Parliamentary Union (IPU) have produced vast corpora of assembly minutes, plenary debates, committee reports, and voting records. Although lots of minutes are available openly on the Web as scanned images or PDFs with document-level metadata for humans to read and study, the minutes are not available as FAIR data with detailed semantic annotations for data-analyses and visualizations that are needed for Digital Humanities research [2], for application development, and for AI-based knowledge discovery [4].

This paper presents a prototypical case of this as part of the larger “Minutes of Multilateralism” (MoM) initiative⁵: The assembly minutes of the League of Nations (LoN) (1920–1946) offer a rich data source for research. Thanks to LONTAD project [9, 10], these valuable materials [1] have been comprehensively digitized and are accessible online for humans to read, but not available as FAIR semanti data for DH research and application development.

⁵ Minutes of multilateralism project: <https://seco.cs.aalto.fi/projects/minutes/>

2 Contingency Questions and System Requirements

LONTAD made it possible to search and access minutes on a *document level* using limited metadata, such as the date of the assembly meeting. More versatile functionalities are needed for DH research. To specify the requirements for the portal, a set of contingency questions to be solved were set:

- **1. Questions concerning an individual representative.** E.g., In what roles and capacity did Fridtjof Nansen, the Polar explorer, participate in LoN meetings?
- **1.2. Questions concerning a member country and its activities** E.g., What countries have been most active in LoN by the number of given speeches?
- **1.3. Questions about places concerned in LoN meetings** E.g., How did the territorial focus of the League of Nations shift in time?
- **1.4. Questions concerning the resolutions in minutes** E.g., When, by whom, and how was the dispute of Åland resolved by the League of Nations?
- **1.5. Questions concerning the contents of the speeches, e.g., events and topics** E.g., What have the representatives of Japan talked about?

To answer questions like these, the Sampo model [3] for semantic LOD portals⁶ and Sampo-UI tool [6, 8]⁷ was used, adapting and re-using the *ParliamentSampo – Parliament of Finland on the Semantic Web* framework [5] to LoN minutes data. All 14 900 minutes (approximately 27 000 pages) were FAIRified using a new data schema and 12 500 minutes speeches were extracted and annotated. As in ParliamentSampo, the speech data was enriched by creating a biographical KG about the ca 3000 speakers in LoN and by data from related external data sources, especially the Lonsea database⁸, Dodis⁹, Metagrid¹⁰, and Wikidata. The enriched data was published in Zenodo.org¹¹ and in a SPARQL endpoint¹² on top of which the semantic portal LEAGUE OF NATIONS SAMPO¹³ was created and opened. This system is described in detail in [7].

3 Using the Portal for DH Research

For example, consider answering the contingency question: "How did the territorial focus of the League of Nations shift in time?" Based on Sampo-UI, the portal offers faceted search application perspectives for searching the instances of

⁶ Sampo portal series: <https://seco.cs.aalto.fi/applications/sampo/>

⁷ Sampo-UI home: <https://seco.cs.aalto.fi/tools/sampo-ui/>; Github: <https://github.com/SemanticComputing/sampo-ui>

⁸ Lonsea: <https://universe.unibas.ch/projects-collaborations/47560>

⁹ Dodis: <https://www.dodis.ch/>

¹⁰ Metagrid: <https://www.metagrid.ch/>

¹¹ LOD at Zenodo: <https://doi.org/10.5281/zenodo.18867718>

¹² LOD data service and SPARQL endpoint: <https://www.ldf.fi/dataset/lon/>

¹³ LoN Sampo portal: <https://minutes.ldf.fi>



Fig. 1. Heatmap of locations spoken about in 1920-1946 (up) and 1931 (down).

the primary classes in the KG. For example, the ca 12 500 instances of speeches can be searched with 12 facets for filtering them by Search content, Speaker, Country of the speaker, Language, Role of the speaker, Speech time, Referenced people, Referenced location, Referenced organization, Referenced date, and Miscellaneous references. Based on the filter-analyze two-step model of Sampo-UI, the filtered instances can then be analyzed and visualized using tabs that can be selected at the top of the search result. In the case of speeches, there are seven tabs that show speeches on a table, by year, on charts, in a custom speaker view, as a network, on maps, and as a heatmap. Fig. 1 depicts (upper part) the territorial topics (references locations) of all speeches with a clear focus on Geneva and central Europe. In the lower part of the figure, the user has selected from the the Speech time facet value 1931: the territorial focus is now clearly on China, probably due to the Japanese invasion in Manchuria, in Spain probably due to establishing the Second Spanish Republic (1931–1939), and in Mexico for some other reason. In this way, the heatmap (as well as other visualizations and data analyses) is able to point out potentially interesting territorial activities and discussions in time in the LoN data for further study. The interpretation of the visualizations can be then studied by close reading the original speeches on the tab Table that shows the actual speeches, or by using the other tabs.

Bibliography

- [1] Cafiero, F., Cointet, J.P., Mallard, G.: Digital accountability can re-legitimate multilateralism. HAL preprint (2025), <https://hal.science/hal-05396546v1>, preprint, HAL Id: hal-05396546v1. Accessed December 3, 2025
- [2] Drucker, J.: The Digital Humanities Coursebook. An Introduction to Digital Methods for Research and Scholarship. Routledge (2021)
- [3] Hyvönen, E.: Digital humanities on the semantic web: Sampo model and portal series. *Semantic Web* **14**(4), 729–744 (2023). <https://doi.org/10.3233/SW-223034>
- [4] Hyvönen, E.: Using the Semantic Web in Digital Humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery. *Semantic Web* **11**(1), 187–193 (2020). <https://doi.org/10.3233/SW-190386>
- [5] Hyvönen, E., Sinikallio, L., Leskinen, P., Drobac, S., Leal, R., La Mela, M., Tuominen, J., Poikkimäki, H., Rantala, H.: Publishing and using parliamentary linked data on the semantic web: ParliamentSampo system for Parliament of Finland. *Semantic Web* **16**(1) (2025). <https://doi.org/10.3233/SW-243683>
- [6] Ikkala, E., Hyvönen, E., Rantala, H., Koho, M.: Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces. *Semantic Web* **13**(1), 69–84 (2022). <https://doi.org/10.3233/SW-210428>
- [7] Leskinen, P., Hyvönen, E., Lionnet, A., Blukacz-Louisfert, B., Bourneuf, P.E., Rodogno, D., Mallard, G., Cafiero, F.: A linked open data service and semantic portal to study the assembly minutes and prosopography of the league of nations (1920–1946). In: *Proceedings of ESWC 2026, Dubrovnik, Croatia*. Springer-Verlag (May 2026), forth-coming
- [8] Rantala, H., Ahola, A., Ikkala, E., Hyvönen, E.: How to create easily a data analytic semantic portal on top of a SPARQL endpoint: introducing the configurable Sampo-UI framework. In: *VOILA! 2023 Visualization and Interaction for Ontologies, Linked Data and Knowledge Graphs 2023*. vol. 3508. CEUR Workshop Proceedings (2023), <https://ceur-ws.org/Vol-3508/paper3.pdf>
- [9] UN Library & Archives Geneva: LONTAD: Total digital access to the league of nations archives. <https://libraryresources.unog.ch/lontad> (2022), digitization project (2017–2022) providing comprehensive online access to the League of Nations archives
- [10] Wells, C.M.: Total digital access to the league of nations archives: Digitization, digitalization, and analog concerns. In: *Archiving Conference*. vol. 16, pp. 12–16. Society for Imaging Science and Technology (2019)