

Using SAMPOSAMPO data linking service and portal to search and detect conflicting data in a cloud of knowledge graphs (Abstract)

Petri Leskinen^{1,*†}, Annastiina Ahola^{1,†}, Heikki Rantala^{1,†}, Jouni Tuominen^{2,1,†} and Eero Hyvönen^{1,†}

¹*Semantic Computing Research Group (SeCo), Aalto University, Finland*

²*Helsinki Institute for Social Sciences and Humanities (HSSH), University of Helsinki, Finland*

Abstract

Keywords

linked data, digital humanities, entity alignment, data validation, semantic portal, data analysis, knowledge discovery

*SAMPOSAMPO – Connecting Everything to Everything Else*¹ (Hyvönen, Ahola, Leskinen, Rantala, et al. 2025; Hyvönen, Ahola, Leskinen, and Tuominen 2025) is new member in the Sampo series on Linked Open Data (LOD) services and portals (Hyvönen 2022) in use in Finland. In contrast to earlier Sampos², it is a “metasampo” based on a cloud of a dozen other Sampos (BiographySampo, AcademySampo, LetterSampo Finland, etc.) and dozen related external data services on the Web (Wikidata, Geni.com, Getty ULAN, etc.).

This paper focuses on one particular novel use case of SAMPOSAMPO: how to detect automatically conflicting and complementary data about entities in an interlinked cloud of CH KGs Leskinen et al. 2026. With provenience information attached, this is useful 1) for data publishers for checking the quality and possible errors in their data by comparing it with datasets provided by the other publishers, and 2) for data consumers for verifying the quality of data based on multiple primary sources. Our first results show some striking conflicts and disagreements between the primary data sources interlinked in SAMPOSAMPO.

The SAMPOSAMPO knowledge graph is a data linking service with resemblance especially to the *viaf.org*³ service (Hickey and Toves 2014) provided by OCLC. The goal of *viaf.org* is “to lower the cost and increase the utility of library authority files by matching and linking widely-used authority files and making that information available on the Web”. In contrast, our work demonstrates how such a KG and LOD service can be also be used to search and detect when the data about the same entity are conflicting or complementary in different data sources.

More information and publications of SAMPOSAMPO and related works are available on the project homepage⁴.

DHNB 2026: Lost in Abundance: Encounters with the Non-Canonical, March 9–13, 2026, Aarhus, Denmark

*Corresponding author.

[†]These authors contributed equally.

✉ petri.leskinen@aalto.fi (P. Leskinen); annastiina.ahola@aalto.fi (A. Ahola); heikki.rantala@aalto.fi (H. Rantala); jouni.tuominen@helsinki.fi (J. Tuominen); eero.hyvonen@aalto.fi (E. Hyvönen)

🌐 <https://www.aalto.fi/en/people/petri-leskinen> (P. Leskinen); <https://www.aalto.fi/en/people/annastiina-ahola> (A. Ahola); <https://seco.cs.aalto.fi/u/rantala3/> (H. Rantala); <https://seco.cs.aalto.fi/u/jwtuomin/> (J. Tuominen); <https://seco.cs.aalto.fi/u/eahyvone> (E. Hyvönen)

🆔 0000-0003-2327-6942 (P. Leskinen); 0009-0008-6369-4712 (A. Ahola); 0000-0002-4716-6564 (H. Rantala); 0000-0003-4789-5676 (J. Tuominen); 0000-0003-1695-5840 (E. Hyvönen)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 Digital Humanities in the Nordic and Baltic Countries Publications – ISSN: 2704-1441

¹SAMPOSAMPO project homepage: <https://seco.cs.aalto.fi/projects/ss/>

²Sampo series of over 20 LOD services and CH portals: <https://seco.cs.aalto.fi/applications/sampo/>

³Virtual International Authority Files: <https://viaf.org>

⁴SampoSampo project homepage: <https://seco.cs.aalto.fi/projects/ss/>

References

- Hickey, Thomas B., and Jenny A. Toves. 2014. "Managing Ambiguity In VIAF." *DLib Magazine* 20 (7/8). <https://doi.org/doi:10.1045/july2014-hickey>.
- Hyvönen, Eero. 2022. "Digital Humanities on the Semantic Web: Sampo Model and Portal Series." *Semantic Web* 14 (4): 729–744. <https://doi.org/10.3233/SW-223034>.
- Hyvönen, Eero, Annastiina Ahola, Petri Leskinen, Heikki Rantala, and Jouni Tuominen. 2025. "How to Create a Portal for Digital Humanities Research Using a Linked Open Data Cloud of Cultural Heritage Knowledge Graphs: Case SampoSampo." In *Proceedings of the Second International Workshop of Semantic Digital Humanities (SemDH 2025), co-located with the Extended Semantic Web Conference 2025 (ESWC 2025)*, vol. 4009. CEUR Workshop Proceedings, June. https://ceur-ws.org/Vol-4009/paper_11.pdf.
- Hyvönen, Eero, Annastiina Ahola, Petri Leskinen, and Jouni Tuominen. 2025. "SampoSampo: A Portal for Studying Enriched Data and Semantic Connections on a Cultural Heritage Linked Open Data Cloud." In *The Semantic Web: ESWC 2025 Satellite Events, Portoroz, Slovenia, June 1 - 5, 2025, Proceedings*, 67–74. Springer-Verlag. https://doi.org/10.1007/978-3-031-99554-5_13.
- Leskinen, Petri, Annastiina Ahola, Heikki Rantala, Jouni Tuominen, and Eero Hyvönen. 2026. "Consistency checking in a cloud of interlinked Cultural Heritage knowledge graphs – first results of using the SampoSampo data service and portal." In *Proceedings of DHNB 2026*. Forth-coming. <https://seco.cs.aalto.fi/publications/2025/leskinen-et-al-samposampo-2025.pdf>.