





Project Gallery

Citizen science archaeological finds on the Semantic Web: the FindSampo framework

Eero Hyvönen^{1,2,*}, Heikki Rantala^{1,2}, Esko Ikkala^{1,2}, Mikko Koho^{1,2},
Jouni Tuominen^{1,2}, Babatunde Anafi^{1,2}, Suzie Thomas³ , Anna Wessman³ ,
Eljas Oksanen³, Ville Rohiola⁴, Jutta Kuitunen⁴ & Minna Ryyppö⁴

¹ Department of Computer Science, Aalto University, Finland

² Helsinki Centre for Digital Humanities, University of Helsinki, Finland

³ Department of Cultures, University of Helsinki, Finland

⁴ Finnish Heritage Agency, Helsinki, Finland

* Author for correspondence: ✉ eero.hyvonen@aalto.fi

FindSampo fosters collecting, sharing, publishing and studying archaeological finds discovered by the public. The framework includes the following: a mobile find-reporting system; a semantic portal for researchers, the public and collection managers to use; and a Linked Open Data service for creating custom data analyses and for application developers.

Keywords: Finland, citizen science, metal detecting, archaeology, linked open data, Semantic Web

Introduction

FindSampo develops a prototype framework system for supporting mobile finds data reporting in the field, and for studying archaeological artefacts discovered and reported by the public. While it is unique in responding to the archaeological conditions in Finland, and in providing solutions to its users' needs (Wessman *et al.* 2019), the framework and its implementation are open source and can be replicated for use elsewhere. FindSampo sits within a broader context of digitising Finnish heritage (Hyvönen 2020), the European Public Finds Recording Network (2021) and the ARIADNEPlus (n.d.) infrastructure project.

Challenges

Metal-detectorists' finds can contribute to archaeological knowledge and research. In Finland, however, it has been laborious to access data regarding new metal-detector finds, especially from a researcher perspective, and there is a backlog in the cataloguing process at the Finnish Heritage Agency preventing up-to-date research. Hence, a user-friendly tool for reporting, viewing, browsing and researching metal-detector finds to access high-quality metadata in a timely manner was needed. We adopted a 'citizen science' approach, conducting surveys, interviews and focus groups for future users to express their preferences.

Received: 20 November 2020; Revised: 27 January 2021; Accepted: 9 February 2021

© The Author(s), 2021. Published by Cambridge University Press on behalf of Antiquity Publications Ltd.

FindSampo framework in action

FindSampo's framework consists of three major components:

- *FindSampo Reporter*: a mobile application demonstrator to minimise the reporting work in the field, foster collaborative work with the Finnish Heritage Agency and create a Linked Open Data repository to be used in applications (Hassanzadeh *et al.* 2020).
- *FindSampo Portal*: a research prototype (FindSampo n.d.) of a portal for studying the finds online.
- *FindSampo Data Service*: based on the Linked Data (LD) approach, standards and best practices of World Wide Web Consortium (W3C) (Heath & Bizer 2011). The linked data service (LD n.d.) with its application programming interface can be used for data analysis in digital humanities research with modest programming skills, and by developers for creating new applications of the find data.

FindSampo is based on the 'Sampo model' (Hyvönen 2020) using the "FAIR guiding principles for scientific data management and stewardship" (GoFair n.d.). This model includes three components. A business model for collating, aggregating and publishing heterogeneous, distributed data from different content providers based on a shared ontology infrastructure. An approach to interface design, where the data can be re-used and accessed independently from multiple application perspectives, while the data reside in a single SPARQL (Protocol and RDF Query Language) endpoint (WC3 2015). A two-step model for accessing and analysing the data, where the focus of interest is first filtered out using a faceted semantic search, and then visualised and analysed by ready-to-use Digital Humanities tools of the portal. Implementing user interfaces based on this model is supported by the open source Sampo-UI framework (SeCo n.d.).

In FindSampo archaeological finds can be searched using the faceted search paradigm (Tunkelang 2009), allowing narrowing of the result set by making orthogonal category value selections, such as object type, material, time period and place, based on underlying ontologies (Figure 1). Once a result set of interest has been found, ready-to-use data analytic tools and visualisations can be applied to it with additional contextual information. For example, it is possible to visualise finds on maps at the same time as seeing protected archaeological sites. If the question is about an individual find, its 'home page' can be studied further (Figure 2).

The FindSampo Data Service with its SPARQL endpoint (WC3 n.d.) and data download facility can be used for custom-made analyses. Different software tools can be employed for this. Figure 3 presents a matrix showing probabilities that two types of items of the same era, here Iron Age, are found in the same area, here a municipality.

Infrastructure: data and ontologies

The data used in FindSampo contain initially approximately 3000 finds transformed into LD and published on the LDF.fi platform (Hyvönen *et al.* 2014). This data model describes the

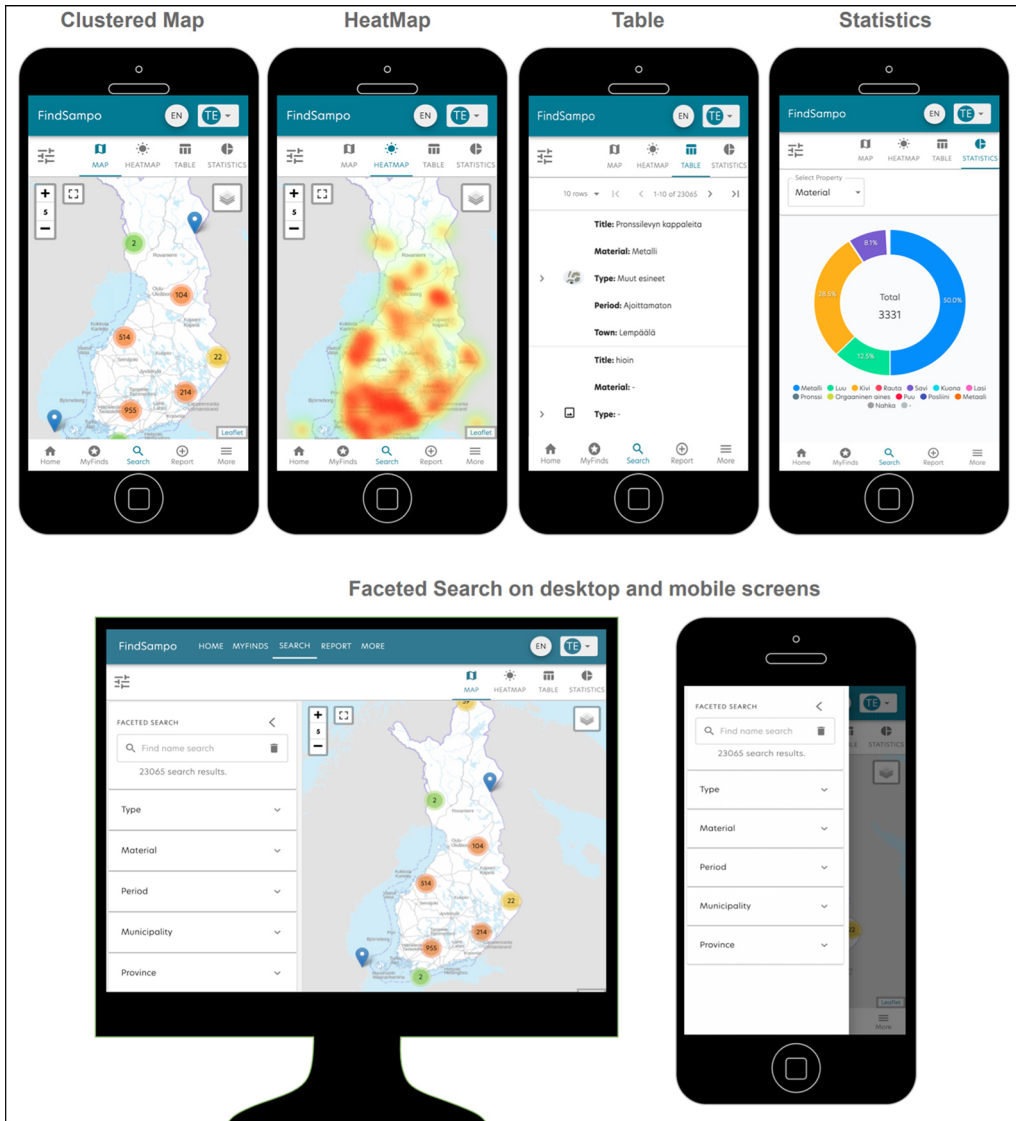


Figure 1. Different views in FindSampo Reporter. From top left: Clusters Map view provides an aggregated view of filtered finds on the map; HeatMap view visualises the filtered finds distribution in colours; Table view lists the finds in a traditional way; Statistics view illustrates statistical distributions of the finds along different facet dimensions, here based on the selected finds' material (graphics by P. Hassanzadeh).

finds in terms of values taken from a set of (hierarchical) ontologies, such as object types, materials and time periods. The ontologies collate heterogeneous data from different data sources and are used to enrich the data by data linking to external data sources and by reasoning based on the Semantic Web logical standards (W3C n.d.). The shared ontology infrastructure includes a new object type ontology of archaeological finds interlinked with the MAO/TAO ontology for Museum Domain and Applied Arts (Finnish Thesaurus and

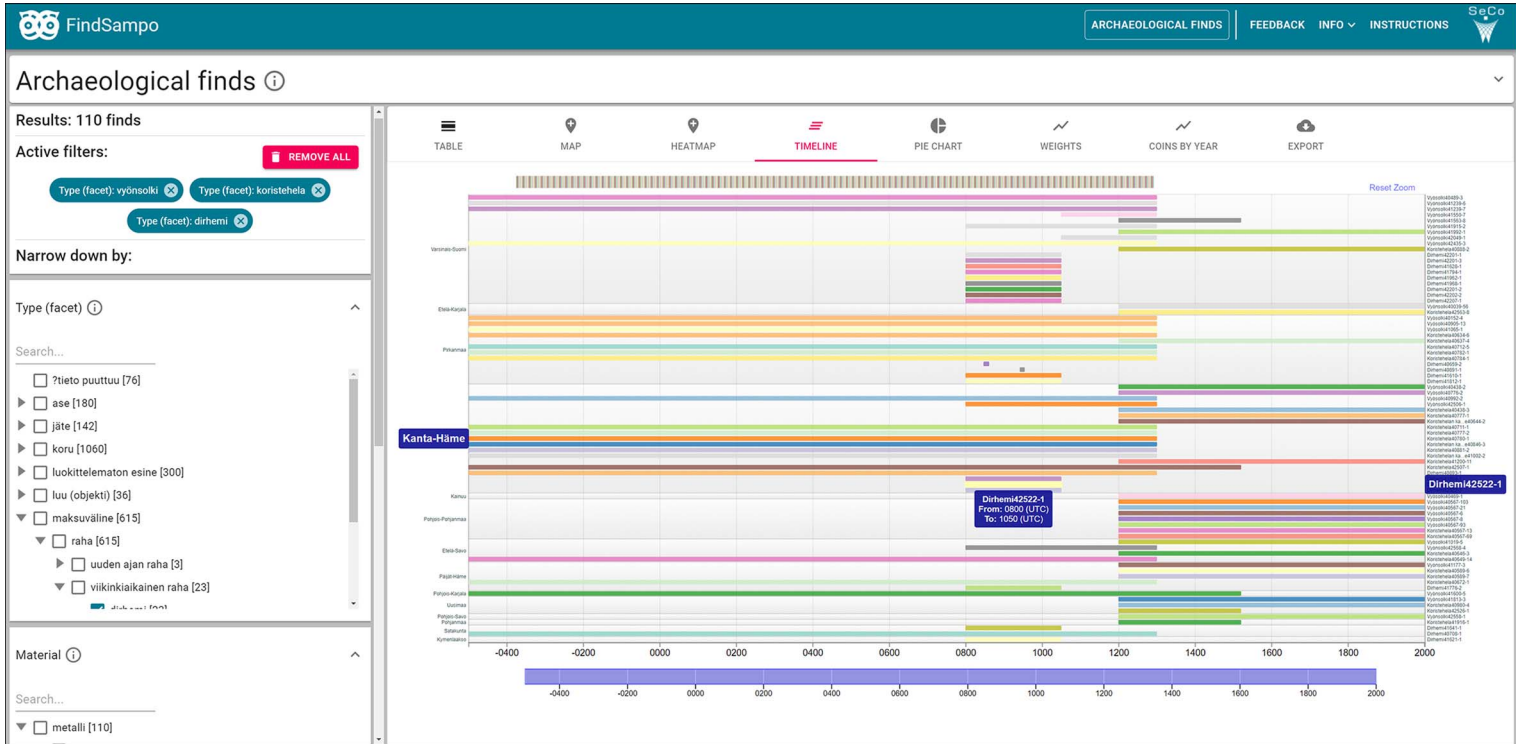


Figure 2. The novel timeline visualisation in the FindSampo portal's user interface (after Anafi et al. 2020). The activated filters are shown on top of the facets in the faceted search section, and the result of the filters is displayed on the activated tab in the results area. The timeline visualisation shows the distribution of the filtered finds over time. Finds are grouped by material type, providing the user with a new perspective on the material distribution of the finds chronologically (graphics by B. Anafi).

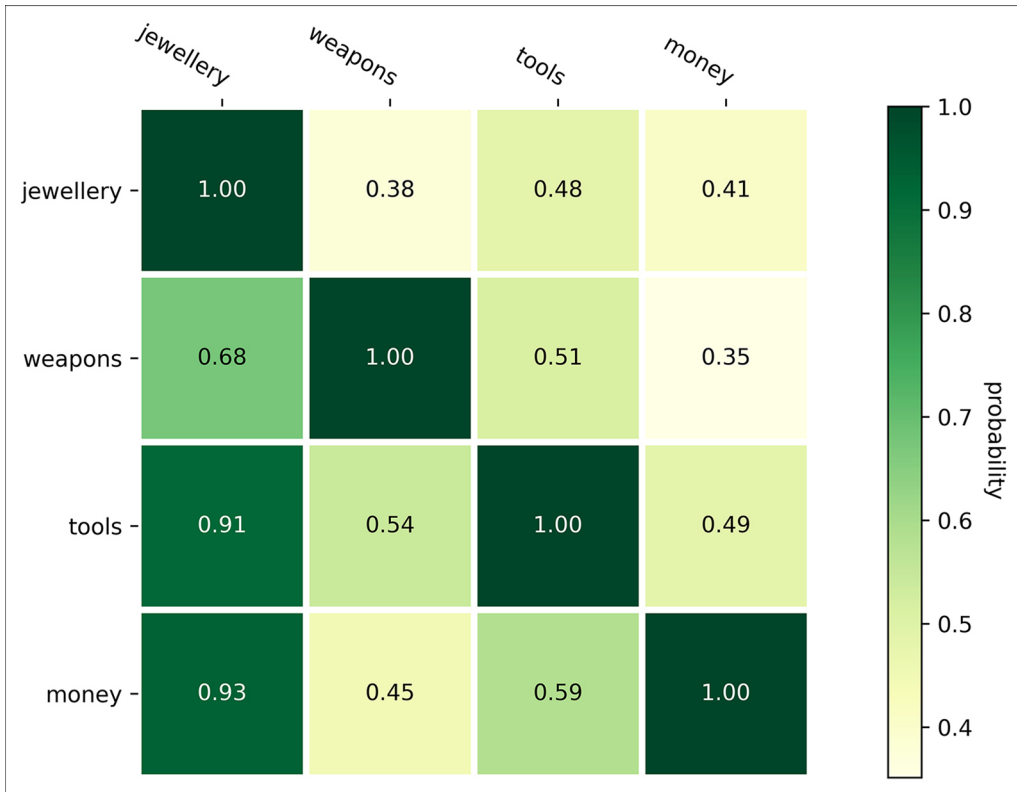


Figure 3. Analysis and visualisation of co-occurring Iron Age object types found in the same municipality, made using Python Matplotlib library and a Google Colab notebook. If coins are found then the probability for jewellery is 0.93, but finding jewellery indicates coins with less probability, i.e. 0.41. Probability for co-occurrence of weapon and coin finds seems low (graphics by H. Rantala).

Ontology Service [n.d.](#)) and the Art and Architecture Thesaurus of the Getty Research Centre (Getty Research Institute [n.d.](#)), and a time-period ontology interlinked with the PeriodO ([n.d.](#)) ontology, as recommended for international semantic interoperability in the ARIAD-NEplus project.

Future visions

The surge of new metal-detected find records in Finland since the 2010s is rewriting our understanding of material culture and associated fields in social, cultural and economic history. To actualise these developments, the FindSampo framework offers novel, groundbreaking qualitative and quantitative research tools to advance digital humanities and citizen science research. Furthermore, a new Marie Skłodowska-Curie project (CORDIS [2021](#)), which began in September 2020, will deploy FindSampo and other Finnish Heritage Agency archaeological data to produce new analysis of large-scale and long-term development of Finnish archaeological landscapes. To test the usability of the FindSampo framework for other find datasets, we plan to apply it to the large Portable Antiquities Scheme ([2021](#))

database managed by the British Museum (2021). These initiatives push towards a deeper understanding of the agency of the public as creators of new knowledge about the past.

Acknowledgements

CSC—IT Center for Science, Finland has provided computational resources for the work.

Funding statement

This article is an output of the research project SuALT—The Finnish Archaeological Finds Recording Linked Open Database (2017–2021), funded by the Academy of Finland (decision numbers 310854, 310859 and 310860). Thanks to AriadnePlus and the Marie Skłodowska-Curie project DeepFIN (grant agreement 896044) for additional funding.

References

- ANAFI, B., M. KOHO & E. HYVÖNEN. 2020. Temporal visualization and data analysis of archaeological finds: case FindSampo. *Conference on Cultural Heritage and New Technologies (CHNT 25)*. Museum Stadt Archäologie Wien, November, 2020. Available at: <https://www.chnt.at/abstract-2020/> (accessed 4 June 2021).
- AriadnePlus. n.d. Available at: <https://ariadne-infrastructure.eu> (accessed 4 June 2021)
- British Museum. 2021. Treasure and the Portable Antiquities Scheme. Available at: <https://www.britishmuseum.org/our-work/national/treasure-and-portable-antiquities-scheme> (accessed 4 June 2021).
- CORDIS. 2021. Assessing archaeological deep time in Finland through spatial exploration 500 BCE–1520 CE. Available at: <https://cordis.europa.eu/project/id/896044> (accessed 4 June 2021).
- European Public Finds Recording Network. 2021. Available at: <https://www.helsinki.fi/en/networks/european-public-finds-recording-network> (accessed 4 June 2021).
- FindSampo. n.d. FindSampo portal. Available at: <https://findsampo.fi> (accessed 4 June 2021)
- Finnish Thesaurus and Ontology Service. n.d. Available at: <https://finto.fi/maotao/fi/> (accessed 4 June 2021).
- Getty Research Institute. n.d. Art & architecture thesaurus® online. Available at: <https://www.getty.edu/research/tools/vocabularies/aat/> (accessed 4 June 2021).
- GoFair. n.d. FAIR principles. Available at: <https://www.go-fair.org/fair-principles> (accessed 4 June 2021).
- HASSANZADEH, P., E. HYVÖNEN, E. IKKALA, J. TUOMINEN, S. THOMAS, A. WESSMAN & V. ROHIOLA. 2020. FindSampo platform for reporting and studying archaeological finds using citizen science, in A. Adamou, E. Daga & A. Meroño-Peñuela (ed.) *Proceedings of the 3rd workshop on humanities in the Semantic Web (WHiSe), Heraklion, Greece, June 2, 2020*: 33–40. CEUR Workshop Proceedings. Available at: <http://ceur-ws.org/Vol-2695/> (accessed 4 June 2021).
- HEATH, T. & C. BIZER. 2011. *Linked Data: evolving the web into a global data space*. Palo Alto: Morgan & Claypool. <https://doi.org/10.2200/S00334ED1V01Y201102WBE001>
- HYVÖNEN, E. 2020. ‘Sampo’ model and semantic portals for digital humanities on the Semantic Web, in S. Reinson, I. Skadiņa, A. Baklāne & J. Daugavietis (ed.) *DHN 2020: digital humanities in the Nordic countries. Proceedings of the digital humanities in the Nordic countries 5th conference, Riga, Latvia, October 21–23, 2020*: 373–78. CEUR Workshop Proceedings. Available at: <http://ceur-ws.org/Vol-2612/> (accessed 4 June 2021).
- HYVÖNEN, E., J. TUOMINEN, M. ALONEN & E. MÄKELÄ. 2014. Linked Data Finland: a 7-star model and platform for publishing and re-using

- linked datasets, in V. Presutti, E. Blomqvist, R. Troncy, H. Sack, I. Papadakis & A. Tordai (ed.) *The Semantic Web: ESWC 2014 satellite events. ESWC 2014*: 226–30. Cham: Springer. https://doi.org/10.1007/978-3-319-11955-7_24
- LD. n.d. FindSampo linked data source. Available at: <https://ldf.fi/dataset/findsampo> (accessed 4 June 2021).
- PeriodO. n.d. Available at: <https://perio.do/en/> (accessed 4 June 2021).
- Portable Antiquities Scheme. 2021. Available at: <https://finds.org.uk> (accessed 4 June 2021).
- Semantic Computing Research Group (SeCo). n.d. Available at: <https://seco.cs.aalto.fi/tools/sampo-ui/> (accessed 4 June 2021).
- TUNKELANG, D. 2009. *Faceted search: synthesis lectures on information concepts, retrieval, and services*. Palo Alto: Morgan & Claypool. <https://doi.org/10.2200/S00190ED1V01Y200904ICR005>
- WC3. 2015. Semantic Web. Available at: <https://www.w3.org/standards/semanticweb/> (accessed 4 June 2021).
- WESSMAN, A. *et al.* 2019. Citizen science in archaeology: developing a collaborative web service for archaeological finds in Finland, in J. Jameson & S. Musteață (ed.) *Transforming heritage practice in the 21st century: one world archaeology*: 337–52. Cham: Springer. https://doi.org/10.1007/978-3-030-14327-5_23