OperaSampo – Opera and Music Theatre Performances in Finland 1830–1960 on the Semantic Web

Eero Hyvönen[1,2], Annastiina Ahola[1], Heikki Rantala[1] and Anne Kauppala[3]

[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*
[3]*Sibelius Academy, University of the Arts, Finland*

Abstract

This paper demonstrates using OPERASAMPO, a LOD service and semantic portal for Finnish opera and music theatre performances in 1830–1960. The novelty of OPERASAMPO lays on its focus on using and studying historical data about the musical performances and persons involved in different roles. OPERASAMPO PORTAL offers faceted search and integrated data-analytic tools for easily searching and analyzing the data, improving the usability of the data for research purposes. The system, data, and software used are openly available on the Semantic Web.

1. Introduction

Reprises[1] is a database of the Sibelius Academy, Finland, that includes evening-specific information about opera, operetta, vaudeville, and other forms of music theatre performances in Finland during 1830–1960. The database covers the performances of the Esplanade Theater in Helsinki (1827–1857), the Swedish Theatre in Helsinki (1860–1961), the Russian Theatre in Helsinki (1868–1917), the Finnish Opera Company (ca. 1873–1879), and music theatre performances in the Turku Play-house (1839–1897). Reprises has been in use for both research and educational purposes at the University of the Arts, Finland, as well as in the music media sector. The data currently includes information about 675 operas, 9075 performances, and 3525 persons involved in different roles (e.g., as singers, conductors, or composers). The main primary sources of the data have been theatre posters, newspaper advertisements, and repertoire books [1, 2, 3, 4, 5].

When using the legacy Reprises portal, performances of the database have been searched by using a traditional search form that specifies values for the Composer, Opera, Singer, Conductor, Keywords, and Time interval of the event. By clicking on an item in the search result list, metadata about the performance (Season, Conductor, Orchestra, Producer, Tickets, Additional information, and Roles) and other performances of the same opera can be studied using close-reading. The operative languages of Reprises are Finnish and English. The languages of the

[1]Available at: http://reprises.uniarts.fi/en/ (no longer accessible on up-to-date modern browsers)

performances are mainly Finnish, Swedish or Russian, and have been indicated in the respective records. The Reprises user interface, however, is no longer accessible on up-to-date modern browsers due to running on older, hard-to-update legacy software. This was to be expected and was one of the reasons for looking for a new solution for browsing and searching the Reprises data.

This paper presents and demonstrates a new Linked Data (LD) approach and implementation, OperaSampo, for searching, browsing, and analyzing the Reprises data. OperaSampo is based on the Sampo Model [6] and Sampo-UI framework, [7], where the data can be filtered using faceted semantic search in multiple application perspectives, and then be analyzed using seamlessly integrated data-analytic visualizations and tools for Digital Humanities research.

2. Ontology and Data

Reprises data was available as a data dump of 16 CSV tables corresponding to class concepts, such as Performances, People, and Roles. The tables were used in a data-driven way for creating an ontology for musical performances: Nine of the tables were transformed into classes and sets of their instances with properties corresponding to table columns. The rest of the tables were transformed into additional properties and sets of property values for the created classes. The created Dublin Core-like ontological data model is introduced in more detail in [8]. The resulting knowledge graph was then published and documented on the Linked Data Finland LDF.fi platform[2]. The data is being enriched from external data sources, such as Wikidata and BiographySampo[3] by mapping local URIs accordingly.

An important aspect in designing the data model was to ensure that the data can be maintained asily by the same people that were in charge of the data in the Reprises database. This way the data could be maintained and edited as native LD instead of having to keep converting the data from CSV tables or other formats each time data needs to be added or corrected. For this purpose, the SPARQL SAHA editor [9] is used.

3. Using the LOD Service and OperaSampo Portal

OperaSampo demonstrates a paradigm shift of using LD and the Sampo model for studying musical performances in Digital Humanities by offering plenty of novel possibilities for retrieving focused information about the contents of the database and analyzing them. For instance, the performers and their performances are of particular interest for the researchers using the OperaSampo data. The OperaSampo Portal can, for example, easily visualize singer's roles as well as their succession and frequency in the career as shown in Fig. 1 for the Finnish opera singer August Alfred Aspegrén.

The portal also lends itself to exploring repertoires of opera houses, or composers' success in opera houses' programs. For example, Fig. 2 shows how Verdi reception in Finland started very slowly and somewhat late. The user has chosen the Performances perspective to the data and can search performances using the seven facets on the left: Composition, Composer, Choir

---

Director, Conductor, Director of Play, Producer, and Opera House. Here Giuseppe Verdi is selected as the Composer. The result set of performances is shown on the right with three tabs available for visualization: Table, Charts, and Timeline. Here the Timeline tab is open. One can see three peaks during three decades (1876, 1886, and 1896), and by using the facets and links one can find out more about what opera houses frequently/rarely performed Verdi, what operas were performed and in what language, the role casts and their possible interconnections, conductors etc.

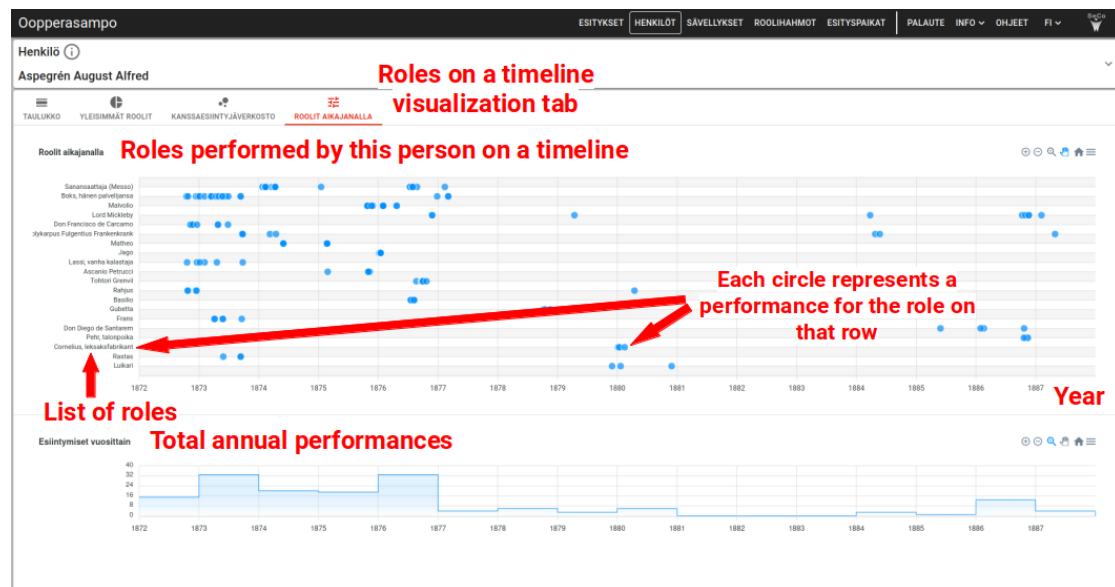Using the OperaSampo portal is demonstrated in more detail in this video[4].

Figure 1: Opera singer August Alfred Aspegrén's roles visualized on a timeline

4. Discussion

There are several databases of musical performances available on the Web, including Svenska operans repertoararkiv[5], Dansk Forfatterleksikon[6], The London Stage Database, 1660–1800[7], Archives de l'Opéra Comique[8], Les Archives du Spectacle[9] Staatsoper Dresden database[10], Operatic Productions in the Nertherlands 1886–1995[11], Music in the Second Empire Theatre[12].

---

[4]Video showing how OperaSampo is used: https://vimeo.com/805493196

[5]https://arkivet.operan.se/repertoar/

[6]http://danskforfatterleksikon.dk/1850t/t1850t.htm

[7]https://www.eighteenthcenturydrama.amdigital.co.uk/LondonStage/Database

[8]https://dezede.org/dossiers/archives-opera-comique/data

[9]https://www.lesarchivesduspectacle.net

[10]http://test.performance.slub-dresden.de/projects/staatsoper-dresden

[11]https://brill.com/downloadpdf/journals/rdj/5/2/article-p79_79.pdf

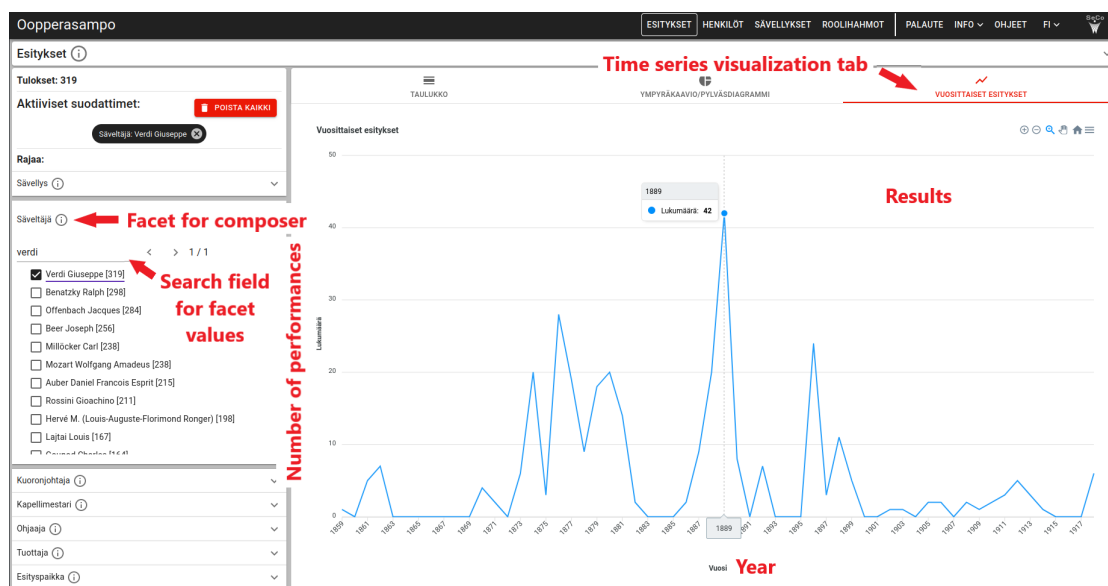[12]http://www.fmc.ac.uk/mitset/index.html#/

Figure 2: Number of performances per year for the compositions of Giuseppe Verdi

However, none of them is based on Linked Data to the best our knowledge.

Linked data has been used for representing and studying musical data before, e.g., in the Linked Jazz[13] system [10], and there are lots of data available for enriching musical data, such as the open music encyclopedia MusicBrainz[14], Wikidata, and online record archives, such as the Live Music Archive[15]. Using LD in representing the relation between performances and scores is discussed in [11], while in [12] LD is used for finding performance recordings in an archive. Ontologies about music, such as the Music Ontology[16], focus on music rather than performances. In contrast, the novelty in OperaSampo lays on its focus on using and studying historical data about the musical performances and persons involved in different roles.

OperaSampo and its source code are openly available online[17]. The underlying LOD service, including a SPARQL endpoint, content negotiation of URIs, linked data browsing, and other services is available at the LDF.fi service[18] [13].

Acknowledgments

---

[13]https://linkedjazz.org/

[14]https://musicbrainz.org/

[15]https://archive.org/details/etree

[16]http://musicontology.com/

[17]Portal: https://oopperasampo.ldf.fi/; source code: https://github.com/SemanticComputing/operasampo-web-app

[18]https://ldf.fi/dataset/operasampo/

References

[1] E. Aspelin-Haapkylä, Suomalaisen Teatterin historia, I–IV, Finnish LIterature Society SKS, Helsinki, 1906–1909.

[2] L. Byckling, Keisarinajan kulisseissa: Helsingin venäläisen teatterin historia 1868–1918, Finnish Literature Society SKS, Helsinki, 2009.

[3] M. Lüchou, Svenska teatern i Helsingfors: repertoar, styrelser och teaterchefer, konstnärlig personal 1860–1975, Stiftelsen för Svenska teatern i Helsingfors/Söderström, Helsinki, 1977.

[4] M. M. van Nieuwkerk, L. Salters, R. M. Helmers, I. Kisjes, Operatic productions in the Netherlands, 1886–1995: from printed annals to searchable performance data, Research Data Journal for the Humanities and Social Sciences 5 (2020) 79–90. URL: https://brill.com/view/journals/rdj/5/2/article-p79_79.xml?language=en.

[5] P. Paavolainen, Arkadian arki: Kaarlo Bergbomin elämä ja työ, II, 1872–1887, Teatterikorkeakoulun julkaisusarja 51, Taideyliopiston Teatterikorkeakoulu, Helsinki, 2016. URL: https://urn.fi/URN:ISBN:978-952-6670-86-7.

[6] 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.

[7] 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.

[8] A. Ahola, E. Hyvönen, H. Rantala, A. Kauppala, Publishing and studying historical opera and music theatre performances on the Semantic Web: case OperaSampo 1830–1960, 2023. URL: https://seco.cs.aalto.fi/publications/2023/ahola-et-al-operasampo-2023.pdf, submitted for peer review.

[9] E. Mäkelä, E. Hyvönen, SPARQL SAHA, a configurable linked data editor and browser as a service, in: The Semantic Web: ESWC 2014 Satellite Events. ESWC 2014, Springer-Verlag, 2014, pp. 434–438. doi:10.1007/978-3-319-11955-7_62.

[10] M. C. Pattuelli, K. Hwang, M. Miller, Accidental discovery, intentional inquiry: Leveraging linked data to uncover the women of jazz, Digital Scholarship in the Humanities 32 (2016) 918–924. doi:10.1093/llc/fqw047.

[11] J. Devaney, H. L. Gauvin, Representing and linking music performance data with score information, in: Proceedings of the 3rd International Workshop on Digital Libraries for Musicology, DLfM 2016, Association for Computing Machinery, New York, NY, USA, 2016, pp. 1–8. doi:10.1145/2970044.2970052.

[12] S. Bechhofer, K. R. Page, D. M. Weigl, G. Fazekas, T. Wilmering, Linked data publication of live music archives and analyses, in: The Semantic Web – ISWC 2017 : 16th International Semantic Web Conference, Springer-Verlag, 2017. doi:10.1007/978-3-319-68204-4_3.

[13] E. Hyvönen, J. Tuominen, M. Alonen, E. Mäkelä, Linked Data Finland: A 7-star model and platform for publishing and re-using linked datasets, in: The Semantic Web: ESWC 2014 Satellite Events, Springer-Verlag, 2014, pp. 226–230. doi:10.1007/978-3-319-11955-7_24.