Using Legislative Change History of Statutes Based on a Linked Open Data Knowledge Graph

Aija Valleala[1], Eero Hyvönen[1,2], Heikki Rantala[1], and Jouni Tuominen[3,2,1]

[1] Semantic Computing Research Group (SeCo), Aalto University, Finland
[2] Helsinki Centre for Digital Humanities (HELDIG), University of Helsinki, Finland
[3] Helsinki Institute for Social Sciences and Humanities (HSSH), University of Helsinki

Abstract. Many tasks related to the enactment, interpretation, and application of laws require examining not only the current version of a statute but also its previous versions. For example, existing Finnish online legal databases provide little support for user-friendly comparison of different versions of the same statute. This paper presents a solution for modelling and visualising statute change histories, enabling easier tracking and interpretation of legislative changes over time. A new prototype application STATUTE HISTORY is introduced to use statutory history data in legal decision making. This application is based on extending an existing Knowledge Graph (KG) (Semantic Finlex) and embedding the prototype into an existing open source User Interface (UI) of a legal web service (LawSampo). Although Finnish legislation is used as a case study, the methods and software presented can arguably also be applied in other countries with a similar legislation system. STATUTE HISTORY enables comparison of different versions of statutes and easy access to related preparatory works. To evaluate the application, a user survey with legal experts was conducted. The respondents found the prototype a clear improvement over existing services, although also identified needs for finer-grained change tracking and more intuitive visual design. User-friendly tracking of statute change history and effortless retrieval of associated preparatory works was deemed essential for legal professionals.

1 Introduction

A central aspect of many tasks related to the enactment, interpretation, and application of laws is the examination of the previous versions of statutes that are currently in force. For example, courts must, as a rule, apply the version of a statute that was in force at the time a crime was committed. If the law has since changed, the correct statute version must be reconstructed. As the statutory language can often be ambiguous or not exhaustive, legal professionals must consult preparatory works, such as government proposals, to understand the legislator's intent and to determine how the statute should be interpreted. Understanding and applying the law in practice often requires access to both historical versions of statues and related preparatory materials. [11]

Much of Finland's legislative information is available online. However, the information is often fragmented across different services and is available in formats that are intended only for a human reader. While several Finnish online legal services provide access to original statutes, amending statutes, and consolidated texts, none of the services currently offer an integrated, user-friendly interface for viewing and comparing

the full change history of individual statute sections. Preparatory works are also available online, but there is no linking between individual statute provisions and the corresponding parts in the preparatory documents. As a result, legal professionals are often required to manually collect and compare documents, sometimes having to engage in what feels like 'detective work' to trace changes to understand the rationale behind the amendments. The lack of interlinking between statute versions and related documents also limits the potential for legal data analytics and advanced information retrieval.

This paper addresses these challenges by presenting a new prototype application STATUTE HISTORY based on an existing open legal KG Semantic Finlex [7] and created by extending the open source legal web application LawSampo [4]. The new application is available online[4], the LOD is in a SPARQL endpoint on the Linked Data Finland platform[5] [5], and the application code was published on Github[6].

2 Related Works

Finlex[7] is a Finnish service that publishes legal information such as statutes, government proposals, case law, and treaties online for public access. The service was completely renewed in 2024 using Akoma Ntoso XML [8] to store statutes, and all data is now published as open data [8] .

Semantic Finlex[9] [7] re-used in our case study was the first initiative to transform Finnish legislation and case law into a machine-readable KG following Semantic Web standards, such as RDF, SPARQL, and the European Legislation Identifier (ELI) and the European Case Law Identifier (ECLI) ontologies. Semantic Finlex provides both SPARQL and REST APIs, enabling programmatic access to legal data for research and application development.

LawSampo[10] [4] system combines legislative data from Semantic Finlex as a simplified KG that is served from SPARQL endpoint, and a web application built on top of the SPARQL endpoint using the Sampo-UI [6] framework. The STATUTE HISTORY prototype UI was created taking the open source LawSampo web application as a starting point and extending it.

There is no single standard or unified approach to how European countries structure and publish their legislative information. In addition to surveying all Finnish legal information services, several publicly available national services from other European countries were reviewed, including those of all Nordic and Baltic countries, to determine how statute history information is provided and visualized. Some services, such

[4] STATUTE HISTORY as part of LawSampo: `https://lakisampo.fi`,`https://lakisampo.fi/fi/statutesHistory/faceted-search/t`

[5] Linked Data Finland LOD: `https://www.ldf.fi/dataset/lawsampo`

[6] Github repository address published after double blind review: `https://github.com/...`

[7] Finlex web service: `https://finlex.fi/en`

[8] Finlex Open data service: `https://www.finlex.fi/en/open-data`

[9] Semantic Finlex: `https://sf.seco.cs.aalto.fi/fi/main`

[10] LawSampo online: `https://lakisampo.fi/en`

as legislation.gov.uk[11] and Restinformation[12], offered a timeline-based view of the history of the statute, but only two of the reviewed services, i.e., Légifance[13], and Riigi Teataja[14], provided a dedicated interface to compare two versions of a statute side by side and visually highlight changes between them.

3 Implementing STATUTE HISTORY

Data Six sample statutes from the new Finlex dataset in Akoma Ntoso format were used in this study.[15] Initial work began with the statute Value Added Tax Act enacted in 1993. However, since available change history extended only as far back as 2005 and missing versions caused challenges with the work, five additional statutes, i.e., Adoption Act, University of Applied Sciences Act, Conscription Act, Narcotics Act, and Universities Act were added to the sample. All five were enacted after 2005 and had been amended multiple times, ensuring a more complete and representative dataset.

The Finlex data lacked references to government proposals for amending statutes. This data was retrieved from the Semantic Finlex using SPARQL. Additionally, some issues in the Finlex data, such as missing or incorrect identifier (eId) attributes, were manually corrected prior to transforming the XML data into RDF, as these errors would have disrupted the subsequent data conversion and integration processes.

Knowledge Graph and Data Model The data transformation in this study resulted in a KG that models the version history of statutes. The graph is expressed in RDF and enriched with structured metadata to represent statutes, their internal components, and their temporal changes. The data model was developed specifically for the statute history view and was not intended to serve as a general-purpose legislative model like the one used in Semantic Finlex. Instead, it was tailored to the particular requirements of version tracking and integration into the LawSampo UI.

The model was constructed as a combination of three sources: the Semantic Finlex data model, the original LawSampo data model, and custom extensions developed in the course of this study. The Semantic Finlex model provided a solid foundation, drawing on established legal ontologies such as the European Legislation Identifier (ELI) and the Semantic Finlex Legislation Ontology (SFL). In contrast, LawSampo's original model supported only the most recent consolidated version of each statute and lacked native support for representing multiple temporal versions. To address this limitation, the model was extended by incorporating relevant classes and properties from the Semantic Finlex model. The model was further extended to incorporate metadata about preparatory works, such as government proposals, which were not fully represented in the original Semantic Finlex data model. A new class, lss:PreliminaryWork, and related properties, such as year, number, and type, were introduced to capture this information. While Semantic Finlex already assigned ELI-compliant URIs to preparatory works,

[11] legislation.gov.uk (UK): `https://www.legislation.gov.uk`

[12] Restinformation (Denmark): `https://www.retsinformation.dk`

[13] Légifrance (France): `https://www.legifrance.gouv.fr`

[14] Riigi Teataja (Estonia): `https://www.riigiteataja.ee`

[15] The statutes were received for research purposes prior to the official release of the renewed Finlex open data, and the data may differ slightly from the currently available version.

additional metadata was needed to meaningfully link them to individual amended provisions within the STATUTE HISTORYiew.

Each statute was represented as a hierarchy of provisions (sections, subsections, paragraphs), where each provision version was modelled as a separate RDF resource identified by a persistent URI. These provision versions were linked to the broader statute, to amendment acts, and to relevant preparatory works using properties such as eli:is_realized_by, eli:amended_by, and eli:related_to. Version-specific metadata, such as enactment dates, entry-into-force dates, and amendment identifiers, was included to support chronological comparison and change tracking. The resulting KG supports flexible SPARQL querying and provides the foundation for visualising statute version histories in the UI.

Data transformation Transformation of Akoma Ntoso XML to RDF was performed using YARRRML [12] and RML [1], two declarative mapping languages designed for Linked Data generation. YARRRML is a human-readable, YAML-based syntax for defining Linked Data generation rules that can be automatically converted into RML, a more verbose RDF-based specification format. The mapping rules were first defined in YARRRML and then converted to RML rules using the yarrrml-parser[16] tool. The resulting RML rules were used as input for the RMLMapper[17], a Java-based processor for executing RDF transformations. The RMLMapper performed the actual data transformation from Akoma Ntoso XML to RDF.

A key advantage of using YARRRML, RML and RMLMapper over other available mapping languages and technologies was the support for user-defined custom functions [2,3]. These functions allowed fine-grained control over how specific values and URI patterns were constructed from the source XML. Several custom functions were developed as part of this study to handle tasks such as trimming strings in a null-safe way, normalising identifiers, and resolving inconsistencies in attribute values. This flexibility was crucial to overcome irregularities in the Akoma Ntoso XML source data and to ensure that the resulting RDF conformed to the target data model.

The resulting RDF data was loaded into the LawSampo triplestore on the Linked Data Finland platform and made accessible via a SPARQL endpoint, which is queried by the LawSampo portal to populate the UI with statute history data.

Extending the LawSampo portal The final step in creating the new STATUTE HISTORY application perspective involved integrating the transformed RDF data into the web application. This required querying the SPARQL API for the statute history data, processing the query results to extract the version history for each section, and rendering the results in the UI.

A SPARQL query was designed to retrieve the necessary data from triplestore. Since LawSampo already included a JSON parser that transforms SPARQL query results into a structured JSON format, particular attention was paid to query design. The query was constructed so that the resulting data would be both compatible with the existing JSON mappings and optimised for further processing in the frontend. Once retrieved, the data was further parsed in the web application application using JavaScript. The logic included identifying all versions of each statute section, sorting them chronologically,

[16] `https://github.com/RMLio/yarrrml-parser`
[17] `https://github.com/RMLio/rmlmapper-java`

and detecting which parts had changed between versions. The parsing process had to account for inconsistencies and edge cases in the underlying Akoma Ntoso XML data, which sometimes affected the reliability of version identification and change detection.

For example, one common issue in the XML data was that a single statute version sometimes included two different versions of the same section. As a result, two conflicting versions of a section could be mistakenly associated with the same statute version. Another recurring problem involved repealed provisions: these lacked machine-readable metadata indicating that the provision had been repealed. Instead, a textual note mentioning that the provision was repealed was added to the section text, while the previous content was retained. A similar issue occurred with sections that were enacted to be in force only for a fixed period. In these cases, the fixed validity period was not represented in metadata but only mentioned in the textual content, in a manner similar to repealed provisions. Some of these issues could be mitigated before the data was shown in the UI while others could not, resulting in certain errors persisting in the displayed content.

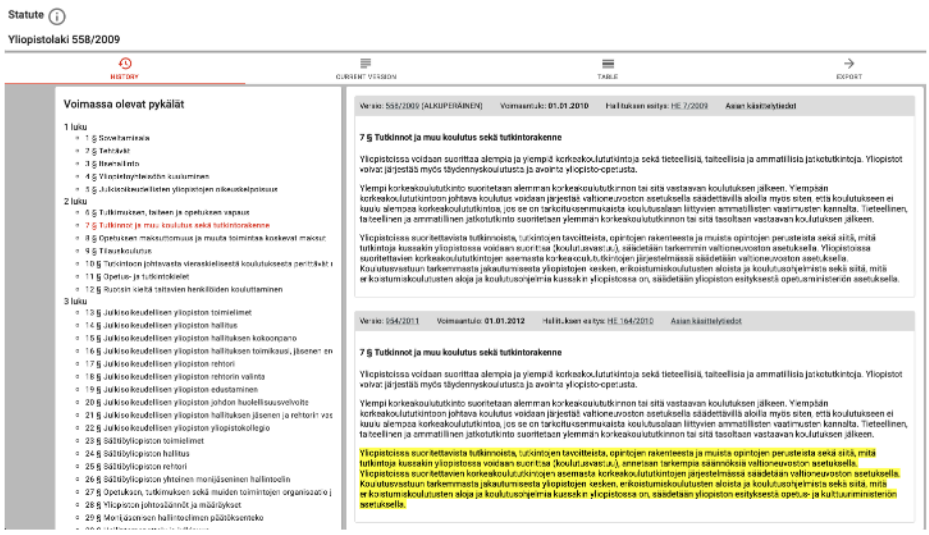

Fig. 1: STATUTE HISTORY view of the Universities Act. A section is selected on the left, and its history is shown on the opened tab on the right. There are also other tabs on the top to view the consolidated current version of the act, metadata about the act and its history, and for investigating the SPARQL query for retrieving the section in the Yasqui editor [9] or for viewing the act in an RDF browser.

The STATUTE HISTORY application was implemented as a new UI component extending the LawSampo portal (cf. Fig. 1). When a user selects a section from a statute's table of contents (on the left), all available versions of that section are displayed on a single page (on the right). Each version includes metadata, such as the date of entry into force, link to full text of the amending statute and links to related preparatory works.

Changes are highlighted visually in yellow to help users distinguish amended content from unchanged text. Highlighting was based on metadata identifying a provision as original version or introduced through an amendment, and when the amendment occurred. The different versions of a provision were iterated in chronological order and each change was highlighted when it first appeared. The highlighting did not rely on textual comparison but on structural and version metadata, which means that even small edits, such as the addition or removal of a few words, could result in the entire provision being marked as changed. Furthermore, this method did not differentiate additions, rewordings, and deletions of text. These limitations remain a point for improvement.

4 User evaluation survey

To evaluate STATUTE HISTORY and how well it fulfils potential users' needs, a user survey was conducted. A total of 15 responses were received from legal professionals working in different sectors of the legal field, including courts, public administration, academia, and private practice. Respondents were asked about their current methods for viewing legislative history, their impressions of the new application, and potential areas of improvement. Overall, the feedback was positive. Users reported that STATUTE HISTORY made it easier to trace changes in legislation and find related preparatory works, which they described as tasks that require substantial manual effort. Respondents especially appreciated having all historical versions of a section visible on a single page, along with direct links to government proposals and other related material. At the same time, several participants highlighted areas for improvement, including the desire for more granular change tracking, such as highlighting only changed parts in the text rather than the entire provision, better indication of which sections have multiple versions, and the ability to compare two versions side by side.

5 Discussion

This paper presented a new solution for visualising the change history of temporally changing statutes using Semantic Web technologies. The Finnish legislation system, based on making modifications to existing legislation, was used in the case study, but the methods presented can also be applied in other countries where a similar legislative system is in use. Building on previous work in Semantic Finlex and LawSampo , STATUTE HISTORY introduces a new UI that enables to view all historical versions of a statute section on a single page, enriched with metadata, such as entry into force dates and links to related preparatory works. The data transformation pipeline was based on Akoma Ntoso XML, YARRRML, and RML, producing LOD that could be queried via SPARQL and served through the LawSampo portal. The approach demonstrated how legal version history, which is often difficult to access in a user-friendly way, can be made more usable and transparent through the application of semantic technologies.

The user evaluation survey indicated that the statute history view is a valuable addition to the toolkit of legal professionals, addressing a clear need for more accessible legislative change tracking. The respondents reported that the view helped streamline what had previously been time-consuming and frustrating work. The ability to view all

historical versions of a statute in one place, and to link directly to related government proposals and other preparatory works, was particularly appreciated. These results support the relevance of structured statute versioning and linked metadata in enhancing legal information systems.

At the same time, several limitations and areas for improvement were identified. One key issue lies in the method of detecting and highlighting changes. The change highlighting logic was based on metadata rather than textual comparison, which meant that small changes in the text often caused entire provisions to be marked as changed. In addition, the system could not distinguish between types of changes and did not offer side-by-side comparison of two different versions. These limitations can hinder the user's ability to understand the precise nature of the changes and may reduce the clarity of the visual representation.

Additionally, inconsistencies in the source Akoma Ntoso XML data, including duplicated sections, missing metadata for repealed provisions, and fixed-term validity expressed only in human-readable notes, presented challenges that could not always be resolved programmatically. While some of these issues were mitigated during data processing, others remain visible in the output. These limitations suggest directions for further refinement in both data modelling and UI design.

Several directions for future research and development emerge from this work. Future research could explore leveraging a wider range of Akoma Ntoso XML elements and attributes, or integrating text comparison algorithms, to support more granular change detection, improve visualisation of modifications, and enable side-by-side comparison of two different versions. The Akoma Ntoso standard allows a much richer set of elements and attributes than was used in the Finlex data. For example, Akoma Ntoso includes elements that can explicitly mark additions and deletions in the text content.

It also allows for indicating validity periods and repeal dates in a structured, machine-readable way, which could help resolve some of the current data quality issues. On the other hand, text comparison algorithms could be used to detect types of changes between different versions of provisions. Furthermore, the Deep Diff algorithm [10] could also offer interesting possibilities, with support for multi-version comparison and cumulative visualisation of change intensity across versions. Finally, the solution could be updated and scaled to include the full Finlex dataset now that it is available.

Acknowledgments

This research was part of the Nordic-Baltic project Achieving the World's Smoothest Cross-Border Mobility and Daily Life Through Digitalisation, funded by the Nordic Council of Ministers and the Cross-border Digital Services (CDBS) programme. Thanks to Anne Kari of the Finnish Digital Agency for collaborations. CSC – IT Center for Science provided computational resources for the project.

References

1. De Meester, B., Heyvaert, P., Delva, T.: RDF Mapping Language (RML): Unofficial Draft, 20 June 2024. Specification Draft, rml.io (2024), `https://rml.io/specs/rml/`, accessed April 2025

2. De Meester, B., Dimou, A., Verborgh, R., Mannens, E.: An ontology to semantically declare and describe functions. In: The Semantic Web: ESWC 2016 Satellite Events. vol. 9989, pp. 46–49. Springer Int Publishing Ag (2016), http://doi.org/10.1007/978-3-319-47602-5_10
3. De Meester, B., Maroy, W., Dimou, A., Verborgh, R., Mannens, E.: RML and FnO: Shaping DBpedia Declaratively. In: Blomqvist, E., Hose, K., Paulheim, H., Ławrynowicz, A., Ciravegna, F., Hartig, O. (eds.) The Semantic Web: ESWC 2017 Satellite Events. pp. 172–177. Springer International Publishing, Cham (2017), https://doi.org/10.1007/978-3-319-70407-4_32
4. Hyvönen, E., Tamper, M., Ikkala, E., Koho, M., Leal, R., Kesäniemi, J., Oksanen, A., Tuominen, J., Hietanen, A.: LawSampo Portal and data service for publishing and using legislation and case law as linked open data on the Semantic Web. In: AI4LEGAL-KGSUM 2022 Artificial Intelligence Technologies for Legal Documents and Knowledge Graph Summarization 2022. pp. 110–114. CEUR Wrokshop Proceedings, Vol. 3257 (2020), https://ceur-ws.org/Vol-3257/paper5.pdf
5. Hyvönen, E., Tuominen, J.: 8-star linked open data model: Extending the 5-star model for better reuse, quality, and trust of data. In: Posters, Demos, Workshops, and Tutorials of the 20th International Conference on Semantic Systems (SEMANTiCS 2024). vol. 3759. CEUR Workshop Proceedings (2024), https://ceur-ws.org/Vol-3759/paper4.pdf
6. Ikkala, E., Hyvönen, E., Rantala, H., Koho, M.: Sampo-UI: A full stack JavaScript framework for developing semantic portal user interfaces. Semantic Web – Interoperability, Usability, Applicability **13**(1), 69–84 (January 2022). https://doi.org/10.3233/SW-210428, https://doi.org/10.3233/SW-210428
7. Oksanen, A., Tuominen, J., Mäkelä, E., Tamper, M., Hietanen, A., Hyvönen, E.: Semantic Finlex: Transforming, publishing, and using Finnish legislation and case law as linked open data on the web. In: Knowledge of the Law in the Big Data Age, pp. 212–228. IOS Press (2019), https://doi.org/10.3233/FAIA190023
8. Palmirani, M., Vitali, F.: Akoma-Ntoso for Legal Documents. In: Sartor, G., Palmirani, M., Francesconi, E., Biasiotti, M.A. (eds.) Legislative XML for the Semantic Web: Principles, Models, Standards for Document Management, pp. 75–100. Springer Netherlands, Dordrecht (2011). https://doi.org/10.1007/978-94-007-1887-6_6, https://doi.org/10.1007/978-94-007-1887-6_6
9. Rietveld, L., Hoekstra, R.: The YASGUI family of SPARQL clients. Semantic Web – Interoperability, Usability, Applicability **8**(3), 373–383 (2017), https://doi.org/10.3233/SW-150197
10. Shannon, R., Quigley, A., Nixon, P.: Deep diffs: visually exploring the history of a document. In: Proceedings of the International Conference on Advanced Visual Interfaces. p. 361–364. AVI '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1842993.1843063, https://doi.org/10.1145/1842993.1843063
11. Valleala, A.: Suomalaisten säädösten muutoshistorian kuvaaminen ja käyttö Lakisampo-järjestelmässä. Master's thesis, Aalto University, Department of Computer Science (October 2024), https://seco.cs.aalto.fi/publications/2024/valleala-msc-2024.pdf
12. Van Assche, D., De Meester, B., Heyvaert, P., Dimou, A.: YARRRML: Unofficial Draft, 26 January 2023. Specification Draft, rml.io (2023), https://rml.io/yarrrml/spec/, accessed April 2025