Analyzing Japanese Law History through Modeling Multi-versioned Entity

Takahiro Komamizu, Yushi Uchida, Yasuhiro Ogawa, Katsuhiko Toyama

Nagoya University, Japan
taka-coma@acm.org, yushi@kl.i.is.nagoya-u.ac.jp,
{yasuhiro, toyama}@is.nagoya-u.ac.jp

Abstract. As law is a blueprint of a society and is changed over time as social environments changed, analyzing histories (change provenances) of laws can reveal important facts such as legislative facts and critical events for the society. Linked Open Data (LOD) has emerged as a preferred method for publishing and sharing open data, however, there is an ontological barrier for publishing law history data as LOD. To break through the barrier, this paper proposes an ontology for law history data of the Japanese statute law. The ontology is inspired from PROV-O and SIOC ontologies. The LOD dataset based on the proposed ontology enables wide variety of analyses on the law history data by simple SPARQL queries. The analyses include simple search, visualization, temporal analysis, data mining, etc. This paper displays parts of the analyses which indicate several legislative facts behind changes of laws. The analyses demonstrate the proposed ontology and LOD dataset are useful for legal data analysis. The proposed ontology is comparable with ELI (European Legislation Identifier) which is designed for EU laws, this paper thus discusses the comparability and future directions of the proposed ontology.

Keywords: Law History, Legal Data Analysis, Multi-versioned Entity

1 Introduction

Linked Open Data (LOD) [2] has become a de facto standard open data publication methodology and has been proliferated over various domains including the legal domain. Laws play a central role in a society and administrative activities are based on the law, therefore, constructing an LOD dataset for the law impacts on the open data movement of not only the legislative and judicial domain but also administrative domains. EU Cases has published a legal LOD dataset of pan-EU law and EU case law, GovTrack offers access to US bills, members of Congress, and so on, and The National Archives, UK provides a legislation API which gives access to law books at various levels for various times. ELI

1 https://lod-cloud.net/
2 http://eucases.eu/
3 https://www.govtrack.us/
4 http://www.legislation.gov.uk/
(European Legislation Identifier) is a promising ontology for EU legislations. Typically, laws are classified into statute laws and common laws, this paper focuses on statute laws (hereafter, laws for short) in Japan. The e-LAWS project in Japan attempted to realize an end-to-end open data platform which handles drafts, deliberations, promulgation, etc. of laws, which realized a support system for drafting partial amendment laws in the Japanese government and a database of currently effective laws. The database is partly opened in e-Gov. In fact, e-Gov (law text database), the Japanese Law Index (law history database) and the Japanese Law Translation Database System (English-translated law text database) are separately operated. Thus, LOD for Japanese laws is highly demanded to realize a universal access to legal documents in Japan.

Laws have been changed according to the changes of societies and environments, and it is necessary to manage histories of laws for enabling analyses on changes of both laws and societies. Since a non-retroactive principle for applying laws and traditional measures for enforcing, amending and repealing laws exist, former versions of laws and repealed laws are necessary for lawyers and administrative officers so that desired versions of laws are easily found [6]. There are also analytical demands for changes of laws like legislative fact discovery in legal studies and social science. Histories for individual laws are currently available online. Though histories are indeed solely useful for drafting laws and knowing amendment tendency, more advanced analyses require efforts for data acquisition, data modeling, and analytical processing. Making the law history data as an LOD dataset is beneficial for universal data accesses on law histories. To this end, designing ontologies which capture the histories is necessary. Although ontologies in the legal domain have been increasingly proposed [3, 7, 8], these are mainly for legislations and documents related to laws.

This paper proposes an ontology design for law histories by utilizing design principles of existing ontologies, PROV-O [5] and SIOC [1]. The basic idea of the ontology is as follows: (1) changes in a history of a law are regarded as version changes of the law, every version of the law is represented by \textit{LawVersion} corresponding with \texttt{prov:Entity}, (2) for access to a law itself rather than versions, the law is regarded as a concept \textit{Law}, (3) when law \(\ell_1\) is changed (i.e., \textit{amended}) by an amendment law \(\ell_2\), \(\ell_2\) is regarded as a version change event which is also represented as \texttt{prov:Activity} and the event is linked with the newly generated version of \(\ell_1\), (4) law version relationships are managed using the SIOC ontology principle (e.g., \texttt{sioc:latest_version} and \texttt{sioc:previous_version}) for easy access to individual versions and sequences of versions.

The proposed ontology design and an LOD dataset based on the ontology enable various analyses for the law history data. Analyzing the original dataset which is represented by HTML tables is laborious on extract-transform-load process for different analyses. While, the ontology enables universal access to the
law history data through a standardized query interface, SPARQL endpoint, and thus makes various analyses easier. This paper reports selected analyses on the LOD dataset, namely, amendment history visualization and classification-based enactment tendency analysis. These analyses showcase that analyses can be performed by simple SPARQL queries, and suggest prospects for more complicated analyses not only on the dataset but also on inter-connected datasets (e.g., DBpedia) with the dataset as a future vision of the dataset.

Contributions of this paper are summarized as follows:

– **ELI-comparable Ontology Design for Japanese Law History**: A main contribution of this paper is to regard a law change as an event and a series of changed laws as versioned entities. The proposed ontology is comparable with European legislation ontology ELI with advantages, thus this paper discusses the comparability and future directions of the proposed ontology.

– **Analytical Use Cases**: Another contribution of this paper is to showcase the usability of the law history LOD dataset for practical analyses and to suggest prospects for more complex analyses. The practical analyses include a simple analysis (i.e., amendment history visualization) and a data mining-based analysis (i.e., classification-based enactment tendency analysis).

The rest of this paper is organized as follows: Section 2 shows an overview of the Japanese law history and the target data. Section 3 introduces the proposed ontology design for the law history data, and Section 4 showcases selected analyses via a SPARQL endpoint. Section 5 discusses the comparability of the proposed ontology with ELI. Finally, Section 6 concludes this paper.

# 2 Overview of Law History in Japanese Statute Law

This paper targets on Japanese statute laws rather than case laws. Since a statute law is a written law that provisions are described as a body of the law, it is necessary to change the body when to change the provisions. A series of changes in the chronological order of a law is called a history of the law.

In Japanese national statute laws, there are six types of laws currently: constitution, act, cabinet order, cabinet office ordinance, ministerial ordinance, and regulation of various governmental organizations. Distinctions among them are jurisdictions. Japanese statute laws are basically enacted through following five steps; (1) drafting law bodies, (2) for acts and cabinet orders, checking the bodies by the legislation bureaus, (3) deliberating the bodies, (4) approving the laws, and (5) promulgating the laws.

The National Diet Library (NDL) services the Japanese Law Index which provides a law history search interface. The interface allows users to search laws by keywords, time periods and categories. Users can observe a law history for each result law. The history is provided as an HTML table for each law, therefore, information extraction techniques [4] can be applied to obtain law histories. In order to obtain all law histories, all laws are searched in the interface. We have
Table 1: Example of consolidations in the Copyright Act (Act No.48, 1970), addition of Article 95-2 by Act No.59, 1984. The first row represents a part of amendment law, the second row shows the original text of a law, and the last row shows the consolidated text of the law, where underlined texts are changed.

<table>
<thead>
<tr>
<th>Amend. Law</th>
<th>The following article shall be added to the next to Article 95 in Section 2 of Chapter 4. (Right of transfer) Article 95-2 (1) A performer has the exclusive right to offer that performer's performance to the public through the transfer of a sound or...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>(Secondary use of commercial phonograms) Article 95 (1) If a broadcaster or cablecaster (hereinafter in this ... … associations under paragraph (5) are provided for by Cabinet Order.</td>
</tr>
<tr>
<td>After</td>
<td>(Secondary use of commercial phonograms) Article 95 (1) If a broadcaster or cablecaster (hereinafter in this ... … associations under paragraph (5) are provided for by Cabinet Order.</td>
</tr>
<tr>
<td></td>
<td>(Right of transfer) Article 95-2 (1) A performer has the exclusive right to offer that performer's performance to the public through the transfer of a sound or...</td>
</tr>
</tbody>
</table>

obtained the data in August 31, 2015. The data include 13,440 laws from February 1886, which is the beginning of the modern legislation systems in Japan, to June 2015 and 61,841 changes. In addition to the law history, NDL provides hierarchical classifications and short titles for laws.

Each law is associated with metadata such as a law number, a title, a promulgation date, an enforcement date, and lapse and repeal dates if exist. The law number consists of a law type (e.g., act and cabinet order), a year and a serial number in the year. For example, the law number of the Copyright Act\(^9\) is “Act No.48, 1970”\(^10\) in Japan. The title is a name of a law. The promulgation date and the enforcement date are dates when a law is publicized and becomes effective. Note that enforcement dates can be different from promulgation dates when enforcements are delayed for dissemination. The lapse date and the repeal date are dates when the law becomes null and void and is repealed.

A law history is a provenance of a law. Changes in the provenance includes enactment, amendment, repeal, suspend, and lapse. Enactment is to make a new law, which is firstly promulgated and enforced later on. Amendment is to change the body (title, provisions, enforcement date, etc.) of a law. Amendment which fully changes the body of a law is called total amendment and that which partially changes the body is called partial amendment. Repeal is to put an end

---


\(^10\) LawID of this act is s45a048.
to a law and the law is no more effective. Suspend is to temporally stop the effect of a law. Lapse is that a law becomes null and void and the law is no more effective but the law still exists.

The most complicated change is the amendment which occurs 268 times per year on average. Amendment changes not only texts of laws but also titles and associated dates. Amendment can change titles of laws so as to make titles suited for the bodies of laws. Amendment can change enforcement dates of laws to delay (or hasten) the effects of the laws. Similarly, if the lapse date of a law is described, the lapse date can also be changed. There are, in general, two ways for amending laws: enlargement and consolidation. The former is to add new provisions to the tail of a former version. The latter is to revise provisions in a former version word by word according to provisions in an amendment law, and, in Japan, the latter is adopted as in many countries. Amendment laws can include descriptions for replacement, addition, deletion and so on [6]. Table 1 showcases examples of consolidations (addition of a statement) on the Copyright Act (Act No.48, 1970), where the first row represents an amendment law, the second row shows the related part of the act and the last row shows the amended text of the act (changed parts are highlighted by underlines). The table shows an article about rights to transfer is added to the next of Article 95 in Section 2 of Chapter 4. Note that amendments can be performed repeatedly, that is a new version of a law can be amended by another amendment law.

Due to the presence of the time lag between the promulgation date and the enforcement date of a law, amendment can performed in between these dates. Therefore, there exist provisions in laws which are promulgated but are never enforced. Such provisions are useful for those who analyze legislation. Figure 1 illustrates two amendment situations. The figure assumes that law $\ell_1$ is enacted and is amended (renamed to $\ell_1'$ for the sake of convenience) by amendment law $\ell_2$. Figure 1(a) represents $\ell_2$ amends $\ell_1$ after $\ell_1$ is enforced. While, Figure 1(b) shows $\ell_1$ amends $\ell_2$ between the promulgation date and enforcement date of $\ell_1$. In the former case, $\ell_1'$ becomes effective when it is enforced and $\ell_1'$ becomes effective when $\ell_2$ is enforced. Note that during the period between the promulgation date and the enforcement date of $\ell_2$, there exists $\ell_1'$ with no effect. In the latter case, $\ell_1$ has never been effective because $\ell_2$ is promulgated before $\ell_1$ is enforced, therefore, there exits a period that the body of $\ell_1$ is known in public even though $\ell_1$ has no effect. Obviously, more complicated amendments can occur (e.g., amendment
A basic idea of the proposed ontology for the law history is that changes of laws are regarded as versioning of laws. Keeping versions in a single graph is beneficial, which enables analysing law history data by simple SPARQL queries, meaning that it does not require to specify graphs corresponding with specific versions. However, it is not convenient if versions of laws separately exist in a graph, due to the large number of changes for querying. For instance, when aggregating the number of laws amending specified two laws, grouping law versions by corresponding laws is required. To cope with this problem, the proposed ontology includes an conceptual class for laws which connect with all versions of individual laws. In the ontology, laws for abstraction are defined as class `law:Law` and versions of laws are defined as class `law:LawVersion`.

Figure 2 illustrates a graphical view of the proposed ontology. According to the previous section, there are six classes (i.e., Constitution, Act, CabinetOrder, etc.) which are subclasses of `law:Law` class. Versioning properties are inspired from PROV-O [5] and SIOC [1] ontologies. PROV-O ontology is designed for data provenance, therefore, it includes generating events (i.e., `prov:Activity`). In the proposed ontology, changing laws are corresponding with the events. That is, the event that a law changes another law is regarded as generation of a new version (`law:LawVersion`) of the changed law. SIOC ontology includes version access properties like `sioc:previous_version` and `sioc:latest_version`. These are useful to access individual versions of laws and sequences of law versions. For the sake of consistency of naming rules, the proposed ontology include properties equivalent to those in the other ontologies with different names and they are marked by `owl:equivalentProperty` to represent the equivalences.

---

11 Namespace `law:` is under discussion, therefore, concrete URIs will be decided in the near future.
In principle, each law change generates a version of the changed law. For instance, promulgation generates a version of the enacted law itself and amendment generates a version of the amended law. In the proposed ontology, an instance of `law:Law` is associated with an instance of `law:LawVersion` which is a version of the law, where the association is represented by `law:hasVersion` for each version and `law:latestVersion` for the latest version. When law $\ell_1$ which partially amends another law $\ell_2$ is promulgated, instances of `law:Law` and `law:LawVersion` for $\ell_1$ are generated, then instances of `law:Law` of $\ell_1$ and $\ell_2$ are connected by `law:partialAmend` and a new instance of `law:LawVersion` for $\ell_2$ is generated by the latest version of $\ell_1$. To maintain the effective periods (discussed in Section 2) and temporal orders of law versions of a law, `law:versionBeginDate` and `law:versionEndDate` are associated with versions, and `law:previousVersion` connects two temporally consecutive versions. The effective period and non-effective period of a law can be calculated using these predicates as follows:

\[
\text{effective\_period} = \begin{cases} 
EFD \text{ to } VED & \text{if } EFD < VED \\
\text{none} & \text{otherwise}
\end{cases}
\]

\[
\text{non\_effective\_period} = \begin{cases} 
VBD \text{ to } EFD & \text{if } EFD < VED \\
VBD \text{ to } VED & \text{otherwise}
\end{cases}
\]

where $VBD$, $VED$ and $EFD$ represent `law:versionBeginDate`, `law:versionEndDate` and `law:enforceDate`, respectively. Note that amendment laws are lapsed just after its promulgation, therefore, these dates are same.

As a result of converting the dataset shown in Section 2 based on the ontology, 571,132 triples are generated. Figure 3 displays a part of the converted data. Due to the space limitation, attributive and systematic properties are omitted. There are three laws, law $h19a128$, law $h24a056$ and law $h30a071$, where the first law is amended by the latter two laws. The following gives an explanation...
of the history in a chronological order. Firstly, law h19a128 is promulgated and law version h19a128_v1 is generated. Secondly, law h24a056 is promulgated and amends h19a128. This amendment is represented by the directed arcs, (h24a056, partialAmend, h19a128) and (h24a056_v1, generate, h19a128_v2). Similarly, law h30a071 amends h19a128. All law versions are associated with individual law:versionBeginDate and law:versionEndDate.

4 Law Amendment Analysis

Law is recognized as a blueprint of a society, analyzing law history data is promising for knowing various aspects of the society. The aspects include states of a society at specific time periods, changes of social blueprints over time, and legislative facts related with law enactments, amendments, etc. The following subsections showcase example analyses of law history data based on the proposed ontology, namely, effective law version detection, amendment history visualization and classification-based enactment tendency analysis. The tail of this section discusses and indicates future analyses with external data sources.

4.1 Effective Law Version Detection

One of main objectives of law amendment analysis is to determine the effective laws at the specified moment. Laws are retroactive to the moment when an activity relevant to laws is done. For a criminal example, suppose that a robber violates the Penal Code at December 10, 2016 and the robber is arrested two years later, the effective version of the Penal Code at December 10, 2016 must be applied to the robber. Therefore, it is necessary to determine a version of a law that should be applied. The effective version of a law can be obtained by the SPARQL query in Listing 1, where <lawURI> and target_date are placeholders for a URI for the law and a specific date, respectively.

Listing 1: SPARQL query for determining an effective version of a law at a given date. <lawURI> and target_date are respectively replaced by a specific law URI and a specific date.
4.2 Amendment History Visualization

A law is a blueprint of a society and the content of the law has been changed over time with respect to the changes of social environments. For example, a Japanese law that minors cannot buy alcohol is enacted because alcohol is considered harmful for the minors. After recognizing an issue that sellers had provided alcohol to the minors without age confirmation, the law is amended so that sellers must confirm ages of buyers with regardless of ages. This example indicates that changes of laws are closely related with social environments and issues.

Analyses with amendment history visualization can indicate social changes in terms of laws. This section include three example visualizations, namely laws about salary of judges (lawID is s23a075) in Figure 4(a), income tax (m32a017) in Figure 4(b), and public assistant (s25a144) in Figure 4(c). The figures illustrate cumulative amendment frequencies over years, where horizontal axes represent years and vertical axes represent cumulative amendment frequencies over years. The frequencies can obtained using the SPARQL query in Listing 2, where <lawURI> specifies a target law.

These examples show that simple visualizations are useful for knowing changes of social environments over time. The first example (Figure 4(a)) represents salaries of judges are constantly changed. This is because the salaries are determined (almost) yearly. The second example (Figure 4(b)) illustrates that the amendment frequency of the law about income tax has drastically changed after World War II (1945). Before the period, the law determines taxes for only people

```
1  SELECT ?lawVersion
2  WHERE {
3    <lawURI> law:hasVersion ?lawVersion.
4    ?lawVersion law:versionBeginDate ?versionBeginDate;
5    law:enforceDate ?enforceDate;
6    law:versionEndDate ?versionEndDate.
7  }
```

Listing 2: SPARQL query for amendment hisotry visualizations. <lawURI> is a specific law URI.
with high income. After the period, the law is changed for applying income tax for all persons, and, due to the increased number of applied persons, amendment frequency has been drastically increased. The last example (Figure 4(c)) is about welfare aids. The example indicates that there are a few amendments on ages for steady economically growth (1973-1991), and the number of amendments has been rapidly increased after the collapse of the bubble economy (1991). This analysis suggests amendment frequencies and economic situations are correlated.

4.3 Classification-based Enactment Analysis

In contrast to the previous analysis, the classification-based enactment analysis indicates attentions to laws of specific topics. As previously mentioned, laws are enacted based on legislative facts. In the proposed LOD, classifications on laws are included (but limited). The classification is hierarchized, for instance, “Construction/CounterDisaster” represents a two-level law classification about counter disaster in terms of construction, where the top level class is “Construction” and the second-level is “CounterDisaster”. With the associated classifications, the analysis in this section can reveal tendency of enactments in terms of the classifications. The query in Listing 3 obtains the number of enacted laws on a specified classification for each year.

Figure 5 displays three classification-based enactment analyses. The first analysis (Figure 5(a)) is of counter disaster. As soon as disaster occurs, laws to support people adversely affected by the disaster are enacted. A notorious disaster in Japan is the Great East Japan Earthquake in 2011, and new five laws are enacted immediately. Similarly, just after the Great Hanshin-Awaji Earthquake occurred in 1995, two laws are enacted in the same year. Figure 5(b) displays enactment frequencies for laws about urban development. After World War II, Japanese cities need restoration works and the government supports their restorations by a municipal enterprise. However, a few years later, the government decided to shrink the enterprise due to the large financial burden. To resist the decision, several cities appeal to the government to support their restoration, and, as a result, 14 cities are selected for restoration supports. Each law is drafted for each of the 14 cities, therefore, 14 laws are enacted during 1949 to 1951. The last example (Figure 5(c)) is about the government bonds in Japan. The special bond act is a one-year effective law and the act is yearly enacted with different name after 1975. However, after 2012 (right-most part of the figure), the act is no more enacted. This is because the act in 2012 has become a three-year effective law.

4.4 Remarks

The analyses in this section introduce the capability of the proposed LOD dataset via simple SPARQL queries and the dataset is potentially useful for finding effective versions of laws, analyzing laws themselves, indicating legislative facts and society situations. Analytical results in this section can be associated with known facts like disasters and economic situations. Since the association is done
Analyzing Japanese Law History through Modeling Multi-versioned Entity

Listing 3: SPARQL query for classification-based enactment analyses. The placeholder is replaced by a specific classification name.

Query:
1. `SELECT year(?date) count(?law)
2. WHERE {
3.   ?law rdf:type law:Act;
4.   ndl:classification "placeholder";
5.   law:promulgateDate ?date.
6. }`
7. `GROUP BY (year(?date))
8. ORDER BY (year(?date))`

Fig. 5: Enactment frequency of laws associated with categories.

(a) Counter disaster.  (b) Urban development.  (c) Government bonds.

5 Discussion: Comparing with ELI

ELI (European Legislation Identifier)\(^2\) is a legislation ontology for EU laws, which can be regarded as a meta-schema of legal documents. ELI abstracts legal documents by three levels (`LegalResource`, `LegalExpression` and `Format`), and ELI provides meta data for each level as well as relationships among these levels. The top-level, `LegalResource`, describes abstract concepts of legal documents, the second level, `LegalExpression`, expresses abstracted contents of the legal documents corresponding with the top level, and `Format` enables to connect the conceptual models of laws with concrete documents like files and URLs. ELI defines several relationships among legal documents such as amendment, consolidation, citation, etc., and it is possible to extend to include dedicated relationships for different legislation systems.

The proposed ontology in this paper and ELI are comparable. The main part of the ontologies are similar in terms of the following two points: (1) `law:Law` is close idea with `LegalResource`, and (2) `law:LawVersion` is similar to `LegalExpression`. The major advantageous differences of the proposed ontology over ELI are versioning-related properties (i.e., `law:versionBeginDate`

and law:versionEndDate), where alternative properties are missing in ELI. The beginning and ending of the versions are when new versions are recognizable. The other but minor differences are the terminology. For instance, ELI does not define a suspend relationship. On the contrary, ELI has lots of desirable representations such as document structures, citation relationships, and links to physical documents. In consequence, the proposed ontology and ELI are comparable with small modification, thus the modification will be the immediate future work.

6 Conclusion

This paper proposes an ontology for law histories in Japan, which can handle multi-versions of laws with change events like amendments. The design of the ontology is based on PROV-O and SIOC ontologies. LOD dataset for Japanese law history based on the proposed ontology enables various analyses in wide range. The analyses shown in this paper illustrates the ability for fundamental analyses (e.g., finding effective versions of laws on specific dates) and advanced analyses (e.g., finding legislative facts based on temporal analyses). These analyses indicate more advanced analyses are expectable. To realize universal ontology design for legal documents, the proposed ontology in compared with a promising ontology, ELI, and the comparison realizes that the interchangeability of the ontologies is an important discussion.

Acknowledgements

This work was partly supported by JSPS KAKENHI Grant Number JP18H03492.

References