

法令沿革 LOD 構築のための DBpedia における法令エンティティの同定

Identifying Legal Entities in DBpedia for Statute History LOD

駒水 孝裕* 小川 泰弘 外山 勝彦

Takahiro Komamizu, Yasuhiro Ogawa, Katsuhiko Toyama

名古屋大学
Nagoya University

Abstract: 本稿では、法令沿革オントロジーの設計および DBpedia 中の法令エンティティへの紐付けについて述べる。社会のさまざまな事柄が法令に関係しているにもかかわらず、日本の法令に関するデータのオープン化はほとんど進んでいない。特に、再利用性の高い LOD (Linked Open Data) としてのオープンデータはほとんどない。本稿で提案する法令沿革オントロジーは次の 2 点を達成するために設計したオントロジーである。(1) 日本の法令を網羅できる。(2) 任意の時点で有効な法令 (法令バージョン) を特定できる。前者は、法令に関するオープンデータの LOD としての公開を促進することがねらいである。後者は、法令における不遡及の原則や経過措置などの理由から必要である。法令の改正は、その変更内容を記述する法令 (改正法令) を施行することにより実現される。このような法令の改正履歴を法令の沿革と呼ぶ。本稿では、法律を対象に設計した先行研究のオントロジー [2] を拡張し、法律以外の法令の沿革も扱えるようにした。拡張したオントロジーをもとに、国立国会図書館・日本法令索引からデータを取得し、法令沿革 LOD を構築した。構築した LOD は、3,412,748 個のトリプルで構成され、106,341 法令を含む。外部ドメインの LOD と接続するために、DBpedia 内の法令エンティティとの紐付けを行った。法令名を用いた単純な紐付けにより、99% の適合率、96% の再現率が達成可能であることを明らかにした。

1 はじめに

法令は社会の設計図であり、社会におけるさまざまな事柄に影響を与える。本稿において、法令とは、憲法、法律、条約、政令、勅令、閣令、府省令、太政官布告を指す¹。我が国では、法令は成文法 (statute law) として運用されている。つまり、法令は文字によって記述され、国民は法令の内容を知ることができる。

法令に関するデータはオープン化が徐々に進められている。例えば、衆議院の制定法律²には制定時点の法律が HTML 形式で公開されていたり、e-Gov 法令検索³では、法令の最新バージョン⁴を検索、XML 形式でダウンロードできる。国立国会図書館・日本法令索引⁵は、法令に関するメタデータ (法令番号、公布年

月日、通称など) と法令の改正履歴である法令の沿革を HTML 形式で公開している。

法令のオープンデータ化は進行しつつあるが、LOD (Linked Open Data) [1] として公開されているものはほとんどない。我が国の法令データで LOD となっているものは DBpedia⁶ と日本の法令 LOD⁷ である。しかし、DBpedia には社会的関心の強い法令のみが収録されており、網羅性は高くない。これは、DBpedia が Wikipedia 記事を LOD 化した経緯から当然と言える。DBpedia における法令の網羅率を把握することは容易でないが、4 節で述べる方法に基づくと、網羅率はおよそ 2% である。一方、日本の法令 LOD は、e-Gov に掲載されている法令を LOD 化したものである。e-Gov は廃止された法令や改正前の法令、一部改正法令をカバーしないため、網羅率はおよそ 8% である⁸。このことから、上記 LOD は法令の網羅率が十分に高くなく、法令に関係するデータ (例えば、判決文やニュース記事) を DBpedia の法令と関連付ける際に、当該法令が

*taka-coma@acm.org

¹告示・訓令や規則 (会計検査院規則や最高裁判所規則など) を法令に含めることもあるが、本稿ではこれらは除外している。

²http://www.shugiin.go.jp/internet/itdb_housei.nsf/html/housei/menu.htm

³https://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/lsg0100/

⁴厳密には、一部改正法令の内容が反映された被改正法令のデータで、各府省が確認したもの

⁵<https://hourei.ndl.go.jp/>

⁶<http://dbpedia.org/>

⁷<https://github.com/lod4all/e-laws-lod>

⁸2020 年 5 月末時点で、8,401 法令を収録 (法令 API (https://www.e-gov.go.jp/elaws/interface_api/index.html) より)

エンティティとして DBpedia に存在しない可能性がある。このような低い網羅性は、法令に関するデータ統合・検索・分析など今後の法令オープンデータ利活用において支障となる。そのため、法令を網羅的に収録した LOD が必要である。

法令は社会の変化に合わせてその内容が改正 (amendment) される。例えば、「新型インフルエンザ等対策特別措置法」(平成 24 年法律第 31 号) は新型コロナウイルス感染症 (COVID-19) に対応するために改正された。法令の改正は、その変更内容を記述する法令 (改正法令) を施行することにより実現される。例えば、「新型インフルエンザ等対策特別措置法」は「新型インフルエンザ等対策特別措置法の一部を改正する法律」(令和 2 年法律第 4 号) によって改正された。このように、法令の内容は社会変化に合わせて書き換えられてきた。法令の改正による不当な不都合・不利益を防ぐため、不遑及の原則や経過措置の制定がある。前者は、法令施行以前の事柄に対して当該法令を適用しないための原則である。後者は、法令施行により既得権益が損なわれる場合に、一時的に法令適用の例外を認める措置である。これらのことから、法令の適用には、その時点で有効な法令の内容を把握する必要がある。

本稿では、日本法令に対する網羅性と法令の沿革をもつ LOD を構築するために、国立国会図書館が提供する日本法令索引に着目する。日本法令索引では、「公文式」(明治 19 年勅令第 1 号) 以降の法令の検索と各法令の沿革の閲覧が可能である。具体的には、法令のメタデータ、沿革、被改正法令、審議経過、外部リソースへのリンクが提供されており、これを LOD 化することで目的を達成する。先行研究では法律を対象として、法令沿革オントロジー [2] の初版を設計した。本稿では、このオントロジーに対して、多種類の法令を記述できるように、法令の種類に対応するクラスおよびプロパティを追加した。具体的には、法律以外の法令 (憲法、条約、政令など) を法令クラスに加え、日本法令索引で用いられている法令の効力の種類を表現するためのプロパティを追加した。

日本法令索引において、法令ページからリンクされる外部リソースは、国立公文書館、総務省、参議院などの政府機関が公開するリソースに限られ、法令本文に関するリンクが中心である。加えて、法律案に関して、審議経過が国会会議録検索システム中の会議録と紐付けられている。法令沿革 LOD では、これら外部リソースへのリンク情報も保存している。一方で、Wikipedia などの政府機関以外のリソースもまた有用である。例えば、Wikipedia では、法令の概説や制定に至った経緯などを集合的にまとめている。このような情報は、官報掲載の「法令のあらまし」や議会提出時の提案理由などからも知ることはできるが、一元化されているという点で Wikipedia の記事は有用である。しかし、集

合的にまとめられている情報は常に正しいとは限らず、一次情報源と併用することによりその信頼性を高めることができる。そこで、本稿では、Wikipedia を元に構成された DBpedia を対象に、法令沿革 LOD 中の法令との Entity Alignment (EA) を行う。

本稿では、ナイーブな EA について議論し、その性能を適合率と再現率の観点から調査する。法令は題名が知られていることが多く、よく使われる通称も法令沿革 LOD に含まれていることから、題名・通称に基づくナイーブな方法で EA が実現可能と考えられる。その際、適合率・再現率の評価のために手動での評価が必要になる。適合率は、提案手法で紐付けされたエンティティが同一かどうかを判定すればよいが、再現率は、DBpedia に含まれるすべての法令エンティティを把握する必要がある。しかし、DBpedia のクラスやカテゴリに日本の法令を網羅するようなものがないため、法令エンティティのすべてを把握することは容易でない。それに対して、本稿では、DBpedia のカテゴリを複数利用する方法と、Wikipedia のテンプレートを利用する方法を適用し、DBpedia 内の法令を近似的に取得した。その結果、題名・通称に基づくナイーブな手法で、適合率 99%、再現率 96% を達成可能であることを示した。加えて、近似的な方法では判別できなかった DBpedia 内の約 500 法令を検出した。本研究の貢献をまとめると以下のとおりである。

- 法令沿革 LOD の構築：日本の法令に対する網羅性と任意時点で有効であった法令のバージョンを管理できるオントロジーを設計した。国立国会図書館・日本法令索引のデータをもとにこのオントロジーを設計し、日本法令索引のデータを LOD 化した。構築した LOD は、3,412,748 個のトリプルで構成され、106,341 法令を含んでおり、日本の法令を最も多く含む LOD である。
- 法令沿革 LOD と DBpedia の法令エンティティの Entity Alignment：法令の題名と通称を使ったナイーブな EA 手法で性能を確認し、適合率 99%、再現率 96% を達成できることを明らかにした。このことから、DBpedia (あるいは Wikipedia) の法令エンティティがそのラベルに題名または通称を利用していることを確認した。また、DBpedia が持つ法令エンティティの法令全体に対する網羅率が 2% 程度であることを明らかにした。

本稿の構成は以下のとおりである。次節では、法令沿革オントロジーについて説明する。3 節では、DBpedia 内のエンティティに対する EA について説明する。4 節では、DBpedia 内に含まれる法令の調査結果を述べる。5 節では、提案オントロジーの発展性について議論し、6 節で本稿をまとめる。

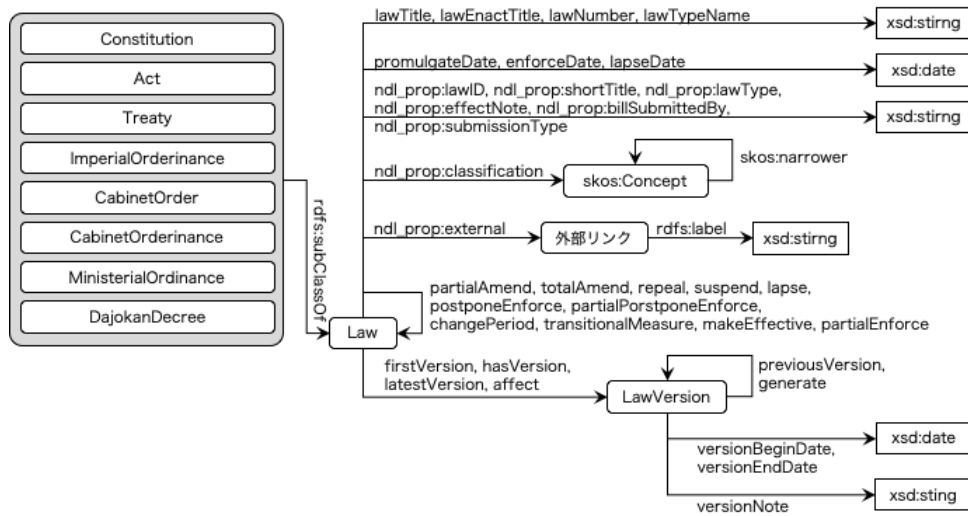


図 1: 法令沿革オントロジー（接頭辞 law は省略）

2 法令沿革オントロジー

本節では、法令沿革 LOD を構築するためのオントロジーについて説明する。まず、法令沿革 LOD の背景知識である法令の改廃 [3] について説明する。その後、オントロジーの概要、法令のメタデータ、法令の改正の表現方法を紹介します。最後に、設計したオントロジーをもとに構築した法令沿革 LOD について述べる。

2.1 法令の内容・効力の変更

法令は社会的変化に合わせて内容や効力が変更される。内容の変更（すなわち改正）は一部改正と全部改正の 2 種類があり、一部改正はある法令中の一部に対する改正であり、全部改正はある法令全体に対する改正である。法令の効力に関する変更には、主に廃止と失効がある。廃止は、法令を無いものにするものであり、廃止された法令は法的効力と文書としての実体が失われる。失効は、法令の法的効力を失わせることであり、廃止とは異なり、文書としての実体は残る。これら以外にも表 3 に示す変更の種類がある。法令の内容や効力に関する変更は、変更の内容を規定した法令を施行することで実現される。

わが国の法令の一部改正では、逐語改正方式が用いられている。この方式では、改正を施す法令中の箇所と改正内容を「改め文」と呼ばれる法令文によって規定し、これを改正前の法令に適用すること（「溶け込み」）により、法令の内容を変更する。例えば、以下に「不正競争防止法等の一部を改正する法律」（平成 30 年法律第 33 号）中の改め文を示す。

第十四条の前の見出しを削り、同条中「前三条」を「第十一条から前条まで」に、「工業標準」を「産業標準」に改め、「廃止に」

の下に「ついて」を加え、同条を第十六条とし、同条の前に見出しとして「（産業標準の確認、改正及び廃止）」を付する。

「第十四条の前の見出しを削り」は文字列の削除、「工業標準」を「産業標準」に改めは文字列の置換、「廃止に」の下に「ついて」を加えは文字列の追加を規定している。「同条を第十六条とし」は、条名を変更することで、文書構造の変更を規定している。このように、改め文は改正前の法令の字句に対する変更および文書構造に対する変更を規定する。

一方で、全部改正を行う改正法令は、「法の全部を改正する。」の一文から始まり、それ以下に新しい法令の内容が記載される。形式的には、改正法令施行時に改正前の法令が効力を失い、改正法令が効力を持つ。全部改正の代表的な例としては、所得税法（明治 32 年法律第 17 号）がある。同法は、これまでに 4 度全部改正されている。これは、税の徴収という制度は継続し、具体的な内容を改めるためとされている。

2.2 法令沿革オントロジーの概要

提案する法令沿革オントロジーを図 1 に示す。このオントロジーでは、表 1 に示す八つの名前空間を使用する。下から四つの名前空間は W3C で勧告されている基本的なオントロジーである。このオントロジーでは、国立国会図書館・日本法令索引における記述を明確に分けるために、語彙を区別して設計した。それぞれの役割は表の「説明」列に記載した。このオントロジーは次の 2 点を達成するために設計されたオントロジーである。(1) 日本の法令を網羅できる。(2) 任意の時点で有効な法令を特定できる。日本法令索引中の法令を網羅的に記述できるようにすることで、(1) を実現する。また、前節で述べた法令の内容や効力の変更を

表 1: 法令沿革オントロジーで用いる名前空間

略称	接頭辞	説明
law	http://www.kl.i.is.nagoya-u.ac.jp/law_history#	本稿で定義する語彙
ndl_jli	https://hourei.ndl.go.jp/#/detail?lawId=	法令エンティティ
ndl_cat	https://hourei.ndl.go.jp/#/result?cc=	日本法令索引の分類
ndl_prop	https://hourei.ndl.go.jp/property#	日本法令索引の定義する関係
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#	RDF の基本語彙
rdfs	http://www.w3.org/2000/01/rdf-schema#	RDF スキーマを記述する基本語彙
skos	http://www.w3.org/2004/02/skos/core#	概念関係を記述する語彙
xsd	http://www.w3.org/2001/XMLSchema#	XML 文書の構造を定義するための語彙

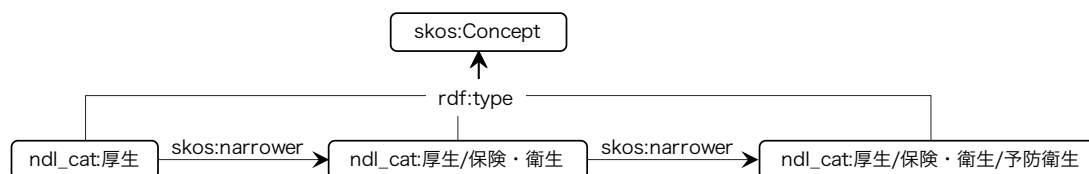


図 2: SKOS を使った概念関係

法令のバージョン変更と捉え、それぞれのバージョンの存在期間を記述することで、(2) を実現する。

このオントロジーにおいて、法令 (Law) と法令バージョン (Law Version) の二つのクラスを定義する。法令クラスは法令番号により識別される法令を表し、法令の新規制定時や改正時ごとの内容は法令バージョンクラスで管理する。法令クラスは、個別の法令種別を表すクラスを子クラスとして持つ。具体的には、憲法 (Constitution)、法律 (Act)、条約 (Treaty)、勅令 (Imperial Ordinance)、政令 (Cabinet Order)、閣令 (Cabinet Ordinance)、府省令 (Ministerial Ordinance)、太政官布告 (Dajokan Decree) をそれぞれ子クラスとした。これらの法令には階層関係があり、上位の法令は下位にある法令に優先し、下位の法令は上位の法令に違反できない。憲法が最上位の法令で、条約は憲法に次ぐ。法律は条約より下位に属し、勅令、閣令、政令、府省令は法律の下位に属する。太政官布告は明治政府初期 (1885 年内閣制度発足以前) の太政官制度時代の法令である。その時点で法律、勅令などという法令の種類はなかった。太政官布告は、その内容の重要性を考慮して、一部は法律扱い、一部は下位法令扱いとみなされた。

2.3 法令のメタデータ

法令のメタデータに関する述語を表 2 に示す。メタデータの記述には、law と ndl_prop の二つの名前空間を用いる。前者は法令についての基本的なメタデータを定義し、後者は日本法令索引のサービスとして付加しているメタデータを定義する。日本法令索引では各法令に階層型分類が与えられている⁹。階層の最上位分類の種類数は、34 個である。分類の例を以下に示す。分類の階層は区切り文字 (/) で表される。

⁹「分類は外部機関から提供を受けたデータを使用しています。」, <https://hourei.ndl.go.jp/help.html#search> (2020/5/20)

- 厚生/保健・衛生/予防衛生
- 租税/特別措置
- 農林/農業/肥料・農薬・植物防疫/植物防疫

分類階層を表現するために、法令沿革オントロジーでは SKOS (Simple Knowledge Organization System)¹⁰ を用いる。図 2 に SKOS で表現した分類階層「厚生/保健・衛生/予防衛生」を表す。それぞれの分類は「概念」を表す skos:Concept クラスのインスタンスとし、分類の上位下位関係を skos:narrower で表現する。

2.4 法令の内容・効力の変更に関する表現

前述のように、法令の内容や効力に関する変更は、その旨を規定した法令を施行することで実現される。本稿では、このような変更を法令のバージョン変更と捉える。例えば、法令の一部改正では一部改正法令を施行し、一部改正法令の内容が被改正法令の最新バージョンの内容に適用される (溶け込む) ことにより、新たな法令バージョンが生成される。この考え方にに基づき、法令沿革オントロジーでは、法令バージョン間にバージョンを生成する関係として記述し、法令クラス間に変更の種類を関係として記述する。「新型インフルエンザ等対策特別措置法」(平成 24 年法律第 31 号) の「新型インフルエンザ等対策特別措置法の一部を改正する法律」(令和 2 年法律第 4 号) による一部改正を例に説明する (簡略化のため、前者を H24A031、後者を R02A004 と呼ぶ)。図 3 は法令沿革オントロジーでの記述を模式的に表したものである。点線枠は法令クラスのインスタンスを表し、実線枠は法令バージョンクラスのインスタンスを表す。法令クラスのインスタンス間に一部改正の関係 (law:partialAmend) を記述すること

¹⁰<http://www.w3.org/2004/02/skos/core#>

表 2: 法令のメタデータ

述語	説明
law:lawTitle	題名
law:lawEnactTitle	制定題名
law:lawNumber	法令番号
law:lawTypeName	法令の種類（憲法，法律，政令，府省令など）
law:promulgateDate	公布日
law:enforceDate	施行日
law:lapseDate	失効日
ndl_prop:lawID	日本法令索引における法令 ID
ndl_prop:shortTitle	日本法令索引に掲載されている通称
ndl_prop:lawType	法令の種類（現行法令，改正法令，廃止法令）
ndl_prop:effectNote	法令の有効性（有効，一部有効，効力なし）
ndl_prop:billSubmittedBy	法律案提出者（内閣「中曽根康弘、外 421 名」など）
ndl_prop:submissionType	法律案提出の種類（閣法，参法，衆法）
ndl_prop:classification	日本法令索引における法令の分類
ndl_prop:external	外部リソースへのリンク

表 3: 法令間の関係を表す述語

法令変更の種類	述語	法令変更の種類	述語
全改	totalAmend	消滅	repeal
改正	partialAmend	(消滅)	repeal
改正附則への改正	partialAmend	適用停止	suspend
失効	lapse	施行停止	suspend
(失効)	lapse	(施行延期)	postponeEnforce
実効性喪失	lapse	(一部施行延期)	partialPostponeEnforce
終了	lapse	(期限変更)	changePeriod
廃止	repeal	(経過措置)	transitionalMeasure
(廃止)	repeal	(一部施行)	partialEnforce
廃棄	repeal	(法的効力付与)	makeEffective

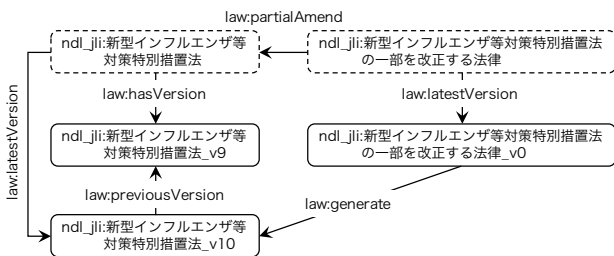


図 3: 法令の内容変更の例（点線枠は法令クラス，実線枠は法令バージョンクラスのインスタンスを示す。）

で，R02A004 が H24A031 を一部改正したことを表現する．R02A004 により一部改正される前の H24A031 の法令バージョンは「ndl_jli:新型インフルエンザ等対策特別措置法_v9」¹¹である．これに R02A004 の施行時の法令バージョンである「ndl_jli:新型インフルエンザ等対策特別措置法の一部を改正する法律_v0」の内容が反映され，H24A031 の新しい法令バージョン「ndl_jli:新型インフルエンザ等対策特別措置法_v10」が生成される．述語 law:generate を用いることにより，これを表現する．

法令の内容や効力に関する変更には種々のものがあ

り，日本法令索引で用いられる種類を表 3 に示す．日本法令索引には，変更の種類について明確な区別が定義されていないものを含まれるため，法令沿革オントロジーでは，一部の種類について以下の解釈のもとで整理し，述語を定義した．

- 括弧付きは，括弧無しと同等
- 終了は，失効と同等
- 消滅 / 廃棄は，廃止と同等
- 上記以外は，適当な英訳から述語を定義

法令と法令バージョンの関係として，法令に対するバージョンであることを law:hasVersion，最初のバージョンを law:firstVersion，最新バージョンを law:latestVersion とそれぞれ定義した．また，法令における法令バージョンの時系列的な関係を表すために，ひとつ前のバージョンを law:previousVersion で示す．

上述のように，一部改正法令の溶け込みは，一部改正法令の内容を表す法令バージョンと，この内容が溶け込んだ被改正法令の法令バージョンの間に law:generate を貼ることで表現する．しかし，現在の日本法令索引の情報だけでは一部改正法令のどのバージョンが溶け込んだのかが自明でない．これは，一部改正法令自体も一部改正される場合があり，さらに施行日の情報が欠

¹¹H24A031 は R02A004 による改正までに 9 回改正されている．

表 4: 法令沿革 LOD の統計

(a) 法令数

(b) 法令間関係の数

法令の種類別	法令数	関係の種類	法令組数
憲法 ¹³	183	一部改正	157,350
法律	13,931	全部改正	2,077
政令	28,010	失効	1,321
勅令	25,359	廃止	14,882
条約	1,497	停止	11
閣令	382	法的効力付与	34
府省令	36,873	施行延期	14
太政官布告・達	136	その他	19
合計	106,371	合計	175,708

けているためのものである。施行日を知るには法令や関連法令（施行令など）の内容を解釈する必要があるため、施行日を系統的に補完することは容易ではない。そのため、施行日が提供されたときのために、法令バージョンを生成する際に関わった法令を `law:affect` で表現することとした。将来、施行日の情報が提供された際に、この `law:affect` を用いて適切な法令バージョン間を述語 `law:generate` で結びつける予定である。

2.5 法令沿革 LOD の構築

法令は日々公布され、改正されるので、日本法令索引のデータは定期的に更新される¹²。今回は、2019年11月に更新されたデータをもとに法令沿革 LOD を構築した。構築した LOD は、3,412,748 個のトリプルで構成され、106,341 法令を含んでいる。この LOD に含まれる種別ごとの法令数を表 4(a) に、法令間関係の種類別の法令組数を表 4(b) に示す。なお、法律でない法令の一部が、例外的に法律と同等として扱われることがある。例えば、ポツダム宣言の受諾により、法律よりも下位の法令に法律と同等に扱われるものがある [3]。日本法令索引では、このような法令を「法律扱い」としており、法令沿革 LOD では、それらに対して、本来の種別に加え、法律も種別として付加している。そのため、表 4(a) では 30 法令が重複して数えられている。

3 DBpedia に対する紐付け

日本法令索引において外部のリソースへのリンクは政府機関が公開する Web ページへのリンクに限られるため、法令沿革 LOD から外部へのリンクもその範囲に限られている。法令沿革 LOD の利用を促進するためには、他のデータとのリンクを構築する必要がある。そこで、本稿ではまず、LOD において中心的な役割を果たす DBpedia¹⁴ へのリンク構築を試みる。具体的に

¹²「基本的に 2 月、5 月、8 月、11 月にデータの更新を行います。」、<https://hourei.ndl.go.jp/help.html>

¹³日本法令索引において憲法と同じ種類として皇室令が含まれる。

¹⁴<http://ja.dbpedia.org/>

表 5: 五つの EA のパターン

パターン	グラフパターン
現行題名 <code>rdfs:label</code>	<code><?s, rdfs:label, 現行題名></code>
現行題名 <code>prop-ja:題名</code>	<code><?s, prop-ja:題名, 現行題名></code>
制定題名 <code>rdfs:label</code>	<code><?s, rdfs:label, 制定題名></code>
通称 <code>rdfs:label</code>	<code><?s, rdfs:label, 通称></code>
通称 <code>prop-ja:通称</code>	<code><?s, prop-ja:通称, 通称></code>

は、法令沿革 LOD の法令エンティティと同一概念である DBpedia 中のエンティティを結びつける Entity Alignment (EA) に取り組む。本稿では、最も単純な EA を実行し、その性能を評価する。

3.1 法令エンティティに対する EA

法令沿革 LOD において法令の特定に寄与する情報としては、法令番号、法令の題名（制定題名、現行題名）および通称がある。例えば「私的独占の禁止及び公正取引の確保に関する法律」（昭和 22 年法律第 54 号）には「独禁法」という通称がある。法令番号はユニークでありもっとも有力な識別子であるが、DBpedia において法令番号に相当する述語は定義されていない。また、関連する述語（`prop-ja:番号`¹⁵）も法令エンティティにおいてほとんど使われていない。一方で、法令の題名や通称は法令を指し示す際に一般的に用いられ、DBpedia における法令エンティティのラベルとして用いられることが期待される。そのため、法令の題名と通称を用いて EA を行う手法を提案する。

DBpedia において、エンティティの名称を示す方法がいくつか存在する。最もよく使われるのは、`rdfs:label` である。DBpedia の日本語版では、日本語での名称を表すために、`prop-ja:題名` という述語を定義している。同様に、エンティティの通称を記述するための `prop-ja:通称` という述語も定義されている。これらが有効な述語であると考え、これらを用いて DBpedia 内の法令エンティティに対する EA を実現する。

提案手法では、表 5 の五つのパターンを用いて EA を行う。各パターンのグラフパターンを含む SPARQL クエリを用いて `?s` のバインディングを取得する。

3.2 適合性の評価基準

DBpedia から取得したエンティティについて、法令沿革 LOD の法令エンティティと同一であるかどうかを人手で評価する。評価基準は適合条件と準適合条件の 2 段階を設定した。適合条件は、法令の識別子である法令番号を基準としている¹⁶。一方で、準適合条件は、法令番号を持たないエンティティについて、エンティティが同じ法令であることが題名などから明らかな場

¹⁵`prop-ja:` <http://ja.dbpedia.org/property/>

¹⁶例外：DBpedia 側で法令番号を間違えている（当該法令番号が明らかに別な法令を指している）場合に内容が正しいと読み取れるとき、適合とする。

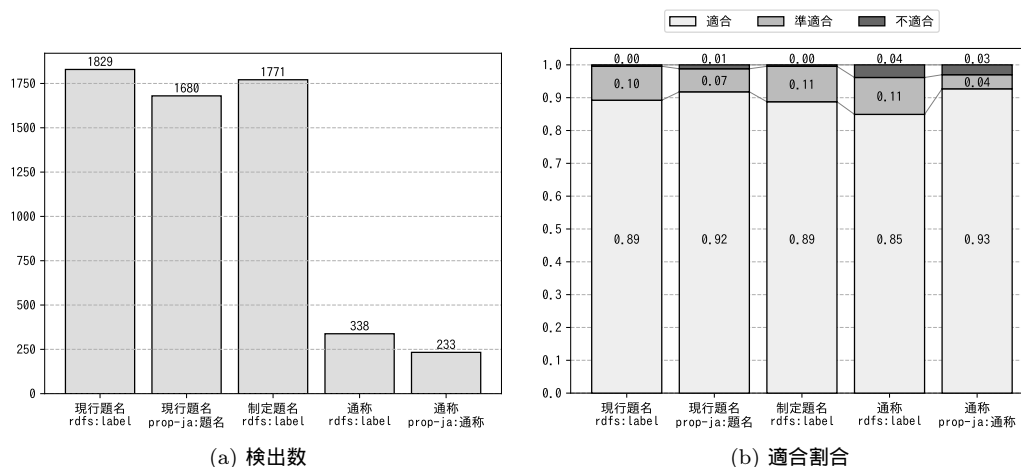


図 4: EA のパターンごとの検出エンティティ数と適合割合

合に、適合とするための条件である。つまり、DBpedia のデータが適切に更新されれば適合条件を満たすエンティティを見落とさないための条件である。

適合条件 以下のいずれかを満たす場合に適合とする。

1. メタデータに法令番号が明記されている。
2. エンティティの Wikipedia の記事に法令番号が明記されている。なお、Wikipedia 記事は prov: wasDerivedFrom から同定する。
3. エンティティがリダイレクトの場合、リダイレクト先に対して 1 または 2 を満たす。
4. 法令が全部改正または廃止・新規制定¹⁷されている場合に、全部改正または廃止・新規制定の前後のいずれかの法令番号について、エンティティが 1 または 2 を満たす。

準適合条件 適合でない場合において、以下のいずれかを満たす場合に準適合とする。

1. エンティティが曖昧性を含む場合に、適合条件を満たすエンティティが曖昧性解消の候補に存在する。具体的には、エンティティが述語 dbo: wikiPageDisambiguates の主語であり、この述語の目的語に適合条件を満たすエンティティが存在する場合に、準適合とする。
2. エンティティのメタデータに法令番号が明記されないが、他のメタデータから同一性が推測できる。同名の別法令を区別するために、公布日の一致を重要な根拠とする。
3. 条約について、法令番号が記載されていないが、当該条約であることが読み取れる。条約は国家間、

国家と国際機構間、国際機構相互間の合意規範であるため、DBpedia (あるいは、Wikipedia) において日本の法令という位置づけで記述されず、法令番号が記載されないことが多い。一方で、その内容や批准した日時から同じ条約であることが読み取ることが可能である。これらから同一の条約であることが読み取れる場合に、準適合とする。

4. エンティティがリダイレクトの場合、リダイレクト先のエンティティが法令自体ではなく、法令が定める制度である。

3.3 結果

EA に用いたパターン全体で、法令沿革 LOD に含まれる 106,341 法令のうち、1,734 法令について適合、223 法令 (適合との重複数 14) について準適合のエンティティをそれぞれ発見できた。図 4(a) と図 4(b) に EA のパターンごとの検出エンティティ数と適合割合をそれぞれ示す。これらの図から、制定題名および現行題名を用いて rdfs:label を使った検出で大多数の法令を検出できていることがわかる。また、どの手法も 85% 以上の適合率である。これは、DBpedia における法令エンティティが、期待通りに、題名や通称などで名前付けされていることを示す。全体を総合すると、適合、準適合、不適合は表 6 のようになる。準適合を正解とすると、全体の適合率は 99% 以上である。

以下でエラー分析を行う。不適合となる理由は、DBpedia に掲載されている法令名が法令沿革 LOD に掲載されていないためである。例えば、「dbr:日本国憲法施行の際現に効力を有する命令の規定の効力等に関する法律の一部を改正する法律」は、DBpedia では昭和 23 年法律第 44 号を指す。実際には、昭和 22 年法律第 244 号も同名の法律である。DBpedia には、後者に相

¹⁷廃止・新規制定も 2.1 節で述べた全部改正と同様に、法令全体の内容を改正するものである。全部改正は既存の制度と新しい制度の継続性を示唆するが、廃止・新規制定はそうでない。

表 6: 法令の種類ごとの適合, 準適合, 不適合の数

法令の種類	適合・準適合	適合	準適合 (適合との重複)	不適合 (適合・準適合との重複)
憲法	12	12	0 (0)	0 (0)
法律	1,400	1,322	86 (8)	24 (5)
条約	185	116	74 (5)	6 (3)
政令	33	22	11 (0)	0 (0)
勅令	134	116	19 (1)	6 (4)
閣令	1	1	0 (0)	0 (0)
府省令	153	122	31 (0)	0 (0)
太政官布告・達	25	23	2 (0)	1 (0)
合計	1,943	1,734	223 (14)	37 (12)

当するエンティティとして「dbr:昭和二十二年法律第七十二号日本国憲法施行の際現に効力を有する命令の規定の効力等に関する法律の一部を改正する法律」がある。この名称は法律案の題名で、制定題名は、昭和23年法律第44号と同名である。ゆえに、後者のエンティティは昭和22年法律第244号と紐付けられない。

また、DBpediaにおいて、通称の使われ方が適合を難しくすることがある。日本法令索引では、「個人情報保護法」という通称は、次の三つの法律の総称である：「個人情報の保護に関する法律（平成15年法律第57号）」、「行政機関の保有する個人情報の保護に関する法律（平成15年法律第58号）」、「独立行政法人等の保有する個人情報の保護に関する法律（平成15年法律第59号）」。しかし、DBpediaでは、「個人情報保護法」は平成15年法律第57号に相当する法令エンティティへのリダイレクトとなっている。そのため、他の2法令の通称でマッチングをすると、「個人情報の保護に関する法律」にリダイレクトされるために、不適合となる。

4 DBpedia 内の法令エンティティ

前節では、単純な方法であっても高適合率で EA が可能であることを示した。一方で、DBpedia に含まれる法令エンティティに対する再現率は明らかでない。再現率を評価するためには、DBpedia 内のすべての法令エンティティを把握する必要がある。しかし、それを正確に取得することは容易ではない。以下では、DBpedia から可能な限り法令エンティティを取得する方法について議論する。その上で、取得した法令エンティティと提案手法の結果から擬似的な再現率（網羅率）を計算し、DBpedia に対する提案手法の評価を行う。

4.1 DBpedia 内の法令エンティティの取得

4.1.1 カテゴリを用いた取得

DBpedia では、法令の種類はクラスとしてではなく、カテゴリとして記述されている。最も関連するカテゴリは、category-ja:日本の法令 である。このカテゴリ

に法令に関するさまざまな下位カテゴリが存在し、法令エンティティは各下位カテゴリに関連付けられる。ただし、下位カテゴリすべてが法令を指すとは限らない。たとえば、「日本の法令」の直下のカテゴリに「地方公共団体の法令」がある。これは、条例と規則を主に指し、本稿の対象ではないため除外する。より煩雑な例として、「日本の法令」の直下カテゴリ「日本の法律」の下に、「刑事訴訟法」というカテゴリがある。このカテゴリに関連付けられるエンティティは、刑事訴訟法に関する事柄一般である。例えば、「dbr:ビデオリンク方式」や「dbr:精神鑑定」などである。

上記の理由から、「日本の法令」の下位カテゴリすべてを利用して、DBpedia に含まれる法令について調べられないことがわかる。そこで、「日本の法令」の下位カテゴリを人手で選別する。具体的には、「日本の法令」以下のカテゴリに関連付けられたエンティティを目視し、そのカテゴリに属するエンティティの大半が法令エンティティかどうかを判断する。具体的には、法令エンティティの割合が90%以上のカテゴリを選出した（表7）。すなわち、法令の種類や各府省の府省令などを表すカテゴリが選出された。

これらのカテゴリのいずれかに関連しているエンティティを取得するために、UNION を用いてグラフパターンを結合した SPARQL クエリを用いた。その結果、1,614 件のエンティティを獲得した。これには、法令の種類を表すエンティティ（例：法務庁令）が一部含まれる。本節の目的は、日本法令索引中の法令が DBpedia に含まれる件数を明らかにすることである。そのため、日本法令索引に含まれない法令や法令でないエンティティを除外する。上記1,614件を目視で確認し、これらを除外した結果、法令エンティティの数は、1,530件となった。以下、この取得方法を「カテゴリ方式」と呼ぶ。

4.1.2 テンプレートを用いた取得

別の方法として、DBpedia が Wikipedia から生成されていることを利用する。Wikipedia には、特定のジャンルの内容を書くためのテンプレートが用意されている。日本の法令を記事として書くためにはテン

表 7: 選ばれた 41 カテゴリ (接頭辞 (category-ja:) は省略)

廃止された日本の法律	日本の講和条約	内務省令	運輸省令
廃止された日本の勅令	日本の基本法	内閣府令	国土交通省令
廃止された府省令	日本の条約	総理庁令	通商産業省令
廃止された政令	日本の法律	総理府令	経済産業省令
日本の旧法令	日本の勅令	文部省令	大蔵省令
日本の旧教育法令	太政官布告・太政官達	文部科学省令	財務省令
日本の旧制教育機関の法令	皇室令	環境省令	法務省令
日本の旧制高等教育機関の法令	閣令	厚生省令	農林省令
府省令_(現存しない種類)	政令	労働省令	農林水産省令
大日本帝国の条約	府省令	厚生労働省令	総務省令
			郵政省令

プレート `template-ja:日本の法令`¹⁸ を用いる。DBpedia において、エンティティの変換元の Wikipedia 記事に使用されているテンプレートは述語 `prop-ja:wikiPageUsesTemplate` で示される。これを利用して、法令エンティティを取得する。具体的には、グラフパターン `< ?s, prop-ja:wikiPageUsesTemplate, template-ja:日本の法令 >` の `?s` へのバインディングを法令エンティティとして取得する。このテンプレートを用いた方法では、1,416 件の法令エンティティを取得した。カテゴリを用いた方法と同様に対象外のエンティティを除外した結果、この方法で取得できた法令エンティティの数は、1,388 件となった。以下、この取得方法を「テンプレート方式」と呼ぶ。

4.2 DBpedia における法令の網羅率

4.1 節で取得した DBpedia の法令エンティティに対して、3 節で提案手法により取得した法令エンティティの網羅率を評価する。網羅率を計算した結果を表 8 に示す。表の (i, j) セルは、 i 方式に対する j 方式の網羅率を表す。方式 i で取得した法令エンティティ集合を L_i とすると、網羅率 $c(i, j)$ は次のように計算する。

$$c(i, j) = \frac{|L_i \cap L_j|}{|L_i|}$$

ただし、 $|\cdot|$ は集合の要素数を表す。

この表から、提案手法によって発見された法令エンティティは、DBpedia に対する各方式に対して 95% 以上を網羅している。一方で、DBpedia に対する各方式は、提案手法に対して 70% 程度しか網羅できていない。このことから、DBpedia 内の多くの法令が、適切なカテゴリに属していなかったり、Wikipedia の日本の法令テンプレートを使用していないことがわかる。DBpedia に対する方式を比べると、テンプレート方式に対するカテゴリ方式の網羅率は 99% であり、カテゴリ方式に対するテンプレート方式の網羅率は 90% である。

この結果から、提案手法では DBpedia に含まれる法令エンティティの大半を網羅していることが明らか

となった。しかし、一部の法令エンティティについて、提案手法は EA に失敗している。提案手法が失敗した要因は大きく三つに分けられる。(1) 法令題名の表記と DBpedia の表記の違い。(2) 日本法令索引における通称の網羅性不足。(3) DBpedia における複数法令をまとめたエンティティの存在。以下でそれぞれの原因について議論する。

(1) 法令題名の表記と DBpedia の表記の違い 新字体・旧字体や英字の混在により法令題名と DBpedia の表記が異なる場合がある。例えば、DBpedia における法令エンティティ「`dbr:屠畜場法`」の「屠」は法令題名では「と」とひらがな表記されている。また、「`dbr:神宮皇學館官制`」は旧字体「學」を含むために、法令沿革 LOD 内の法令エンティティと EA ができなかった。加えて、英字を含む「`dbr:PTA・青少年教育団体救済法`」のような場合、法令題名では全角で記述されるため、提案手法では EA はできない。

(2) 日本法令索引における通称の網羅性不足 日本法令索引に掲載されている通称は一般的な通称を網羅しているわけではない。そのため、DBpedia の法令エンティティが日本法令索引に出現しない通称を用いて表現される場合、EA に失敗する。例えば、「`dbr:駐留軍用地特措法`」の日本法令索引における法令名は、「日本国とアメリカ合衆国との間の相互協力及び安全保障条約第六条に基づく施設及び区域並びに日本国における合衆国軍隊の地位に関する協定の実施に伴う土地等の使用等に関する特別措置法」であり、日本法令索引においてその通称として、「駐留軍用地特別措置法」と「米軍用地特措法」があるが、いずれも DBpedia での呼称「駐留軍用地特措法」とはマッチしない。

(3) 複数法令をまとめたエンティティの存在 DBpedia は Wikipedia 記事から生成されている。Wikipedia の記事はその内容によっては複数の法令をまとめて一つの記事にしていることがある。これにより、DBpedia のエンティティが複数の法令をまとめて一つのエンティティになってしまう。例えば、「`dbr:学校令`」は「帝国大学令」(推定:大正 8 年勅令第 12 号)、「師範学校令」

¹⁸`template-ja: http://ja.dbpedia.org/resource/Template:`

表 8: DBpedia に含まれる法令エンティティと提案手法の網羅率

	提案手法 (1,943)	カテゴリ (1,614)	テンプレート (1,416)
提案手法	1,943	-	0.76 (=1,475/1,943)
カテゴリ	1,530	0.96 (=1,475/1,530)	-
テンプレート	1,388	0.99 (=1,374/1,388)	0.71 (=1,374/1,943)
			0.90 (=1,376/1,530)
			-

(推定: 明治 19 年勅令第 13 号)「小学校令」(推定: 明治 19 年勅令第 14 号)「中学校令」(推定: 明治 32 年勅令第 28 号)をまとめたエンティティである。上記の推定法令番号は、学校令の Wikipedia 記事中の記載とリンクから推定したものである。このように、複数の法令をまとめた場合、提案手法で紐付けるためには、「学校令」が日本法令索引において通称として掲載されている必要がある。しかし、実際にはこのような記載はなく、上記 4 法令は学校令と提案手法では紐付けられない。このような複数法令をまとめたエンティティは、今後も出現する可能性が高い。例えば、`dbr:新型インフルエンザ等対策特別措置法` は、現在「新型インフルエンザ等対策特別措置法」(平成 24 年法律第 31 号)を表している。しかし、新型コロナウイルス感染症に対処するための改正に伴い、元の Wikipedia 記事¹⁹では、一部改正法律「新型インフルエンザ等対策特別措置法の一部を改正する法律」(令和 2 年法律第 4 号)も同記事内に掲載されている。今後 DBpedia に反映される際に、この一部改正法律も含んだエンティティとなる。その際に、平成 24 年法律第 31 号は提案手法により紐付けられるが、令和 2 年法律第 4 号は紐付けられない事態が起こりうる。

5 議論

本稿で提案した法令沿革オントロジーおよび法令沿革 LOD は、日本法令索引に掲載されている法令の沿革を基に構築している。今後、法令沿革 LOD の活用性を向上させるための改善点として大きく分けて以下の 3 点がある。(1) 不足情報の補填、(2) 沿革表現の統一、(3) 通称の拡充。これらについて以下で議論する。

(1) 不足情報の補填: 施行日の欠落 日本法令索引における法令沿革は「公布日」が基準となっている。しかし、実際に法令が効力を持つ時点は「施行日」からである。公布日と施行日が同日である場合もあるが、法令の内容を社会に広める期間として、公布日と施行日の間に周知期間を設ける。特に、社会的影響の大きい法令は、長い周知期間が設けられる。例えば、民法等の一部を改正する法律(令和元年法律第 34 号)は、公布日が令和元年 6 月 14 日で、施行日は令和 2 年 4 月 1 日である。この一部改正法令が被改正法令に適用され、

¹⁹<https://ja.wikipedia.org/wiki/新型インフルエンザ等対策特別措置法>

改正内容が効力を持つのは施行日からである。しかし、日本法令索引には施行日の情報が欠落しており、効力の開始時点を正確に記述できない。

(2) 沿革表現の統一 表 3 に示したように、日本法令索引において沿革の種類は 20 種類あるが、その表記に対する意図が不明確である。本稿では、不明瞭な沿革を一定の基準で解釈し、オントロジーを構築したが、より正確な意味での沿革を作ることが望ましい。

(3) 通称の拡充 法令沿革 LOD をより広く利用可能にするためには、多くのデータと統合利用可能にするべきである。例えば、法令に関する会議録やニュース記事、SNS を収集・統合することで法令に対する社会的背景や国民からの法令に対する反応を分析することが可能になる。しかし、これらの情報源において法令を示す際に呼びやすい言い回しが通称として使われることが多く、題名はあまり使われない。このような言い回しを収集し、通称として掲載することで、法令に関するデータを広く統合し分析することが可能になる。

6 まとめ

本稿では、法令の改正履歴(沿革)を記述するための法令沿革オントロジーを設計し、日本法令索引からこのオントロジーに沿った法令沿革 LOD を構築した。法令沿革 LOD は、日本法令を最も多く有する LOD であり、法令データを統合する際の中核をなすことが期待される。本稿では、DBpedia に対して Entity Alignment を施し、単純な方法で 99% の適合率、96% の再現率を達成できることを示した。また、DBpedia における法令の網羅性が 2% 程度であることを示し、法令沿革 LOD による法令網羅性の重要性を強調した。

謝辞 本研究の一部は JSPS 科研費 JP18H03492 および JP18K18056 の助成を受けたものである。

参考文献

- [1] C. Bizer, T. Heath, and T. Berners-Lee. Linked Data - The Story So Far. *Int. J. Semantic Web Inf. Syst.*, 5(3):1-22, 2009.
- [2] T. Komamizu, Y. Uchida, Y. Ogawa, and K. Toyama. Analyzing Japanese Law History through Modeling Multi-versioned Entity. In *CKG@ISWC 2019*, 2019.
- [3] 法制執務研究会. 新訂 ワークブック法制執務 第 2 版. ぎょうせい, 2018.