

法律分野における大規模引用グラフの構築

丸山拓海¹ 菅原祐太² 西野裕貴³
稲垣有二⁴

法律実務では、法令や判例、専門書籍などの多様な情報源を総合的に検討する必要がある。これらの複雑な参照関係を効率的に把握するため、文書間の関係性を明示的に表現できるグラフデータが注目を集めている。しかしながら、既存の法律分野のグラフデータは、法令や判例を中心としており、実務において重要な参考情報となる専門書籍やガイドラインを含むグラフデータは存在しない。そこで本研究では、日本の法律領域を対象として、法令・判例・書籍・ガイドラインを包括的に結びつける大規模な引用グラフの構築を目指す。これにより、実務における情報収集や分析の効率化を担う、高度な法律実務支援システムの実現に貢献する。

1 はじめに

大規模言語モデル (Large Language Model: LLM) の発展に伴い、検索拡張生成 (Retrieval-Augmented Generation: RAG) [8, 13]の研究が活発に行われている。これは、専門的知識や参考文献を文脈として言語モデルに与えることで、回答の品質を向上させる手法である。近年では、グラフデータを活用した手法も注目を集めている [6, 9, 10, 16, 18]。

グラフデータは、文書間の参照関係や論理的なつながりを明示的に構造化できるため、法律文書のような相互参照が多い分野で特に有効である。実際、グラフデータが整備されている諸外国の法律文書を対象とした研究では、グラフデータベースから得られる知識とLLMを組み合わせることで、テキスト情報のみを用いた質問応答手法と比べて、高い性能を達成できることが報告されている [1, 11, 12]。

日本においても、法令データの整備 [26]や民事判決のオープンデータ化 [24, 25]が進展し、法律分野の文書へのアクセスと活用が容易になりつつある。これに伴い、より実務に即した応用研究の実現に向けて、様々な法律文書の関係性を構造化したグラフデータの整備が期待されている。

しかし、既存の法律分野のグラフデータは法令や判例に主眼が置かれており、実務において重要な参考情報となる専門書籍やガイドラインを含むグラフデータについては、十分な取り組みがな

されていなかった。そこで本研究では、日本の法律領域を対象として、法令・判例・書籍・ガイドラインを包括的に結びつける大規模な引用グラフの構築を目指す。

2 引用グラフ

本研究で構築する引用グラフは、頂点として大きく以下の2種類を扱う：

- ・コンテンツ: 書籍・ガイドライン・判例・法令。
- ・チャンク: コンテンツに含まれるテキストを一定の単位で分割したもの。書籍・ガイドラインは、本文テキストを固定長で分割し、チャンクとした。判例の場合は判決要旨をチャンクとみなし、法令の場合は、条・項・号のそれぞれの単位で分割した。

以下では、各コンテンツの詳細と、それらの引用グラフ上での表現方法について述べる。

2.1 書籍

弁護士ドットコム株式会社が提供する、BUSINESS LAWYERS LIBRARY^{*1}及び弁護士ドットコムLIBRARY^{*2}で閲覧できる、弁護士・企業法務向けの法律書籍・雑誌のうち約1,800件を対象とする。書籍は、書籍頂点と書籍チャンク頂点の2つを用いて表現される (図1左)。

- ・書籍頂点: タイトルや出版日・概要文などのメタ情報を持つ。
- ・書籍チャンク頂点: 一定の長さで本文を分割したテキスト情報とチャンクの順序情報を持つ。

書籍チャンクが書籍の一部であることを表現するため、書籍頂点と書籍チャンク頂点の間を親子関係を表す辺で接続する。また、書籍チャンクから法令・判例への引用関係を表す辺については、書籍チャンクの本文テキストをもとに、ルールベースによる法令名抽出(3章)や判例名抽出(4章)を用いて構築する。

2.2 ガイドライン

「ガイドライン」とは、法令を遵守するための指針が記載された、各省庁が公表している文書を指す。各省庁のWebページから手作業で収集したガイドラインのうち約1万件を対象とする。ガイドラインごとにファイル形式は異なっており、それぞれの形式からテキスト情報のみを抽出する前処理を介した上で頂点を構築する。ガイドラインも、書籍と同様にチャンク単位に分割し、ガイドラインチャンクが持つ本文テキストに対して、法令名抽出・判例名抽出を適用することで、法令・判例への引用関係を表す辺を構築する。

2.3 判例

株式会社LICが提供する判例秘書^{*3}の約20万件の判例を対象とする。判例本文には以下のような情報が記述されている：

- ・判決年月日: 判決・命令・決定が言い渡された年月日。

¹ 正会員 弁護士ドットコム株式会社
ta.maruyama@bengo4.com

² 非会員 弁護士ドットコム株式会社
y.sugawara@bengo4.com

³ 非会員 弁護士ドットコム株式会社
nishino@bengo4.com

⁴ 非会員 弁護士ドットコム株式会社
inagaki@bengo4.com

^{*1} <https://www.businesslawyers.jp/lib>

^{*2} <https://library.bengo4.com/about>

^{*3} <https://www.hanreihisho.com/index.html>

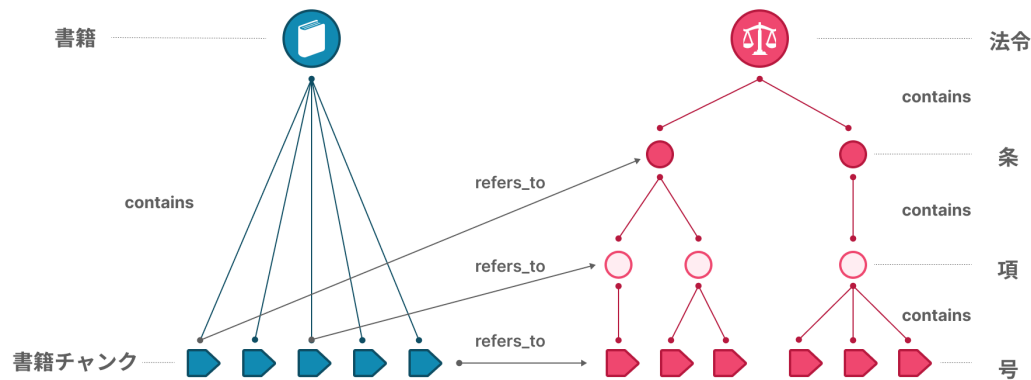


図1 書籍と法令のグラフ構造。(図左)書籍の構造。親子関係の辺(contains)により、書籍頂点と書籍チャンク頂点の関係性を構造化している。書籍から法令や判例への参照関係(refers_to)は、書籍チャンク頂点を持つテキスト情報に基づいて構築する。(図右)法令の構造。条・項・号をそれぞれ頂点とし、木構造で表現する。法令から法令への参照関係(refers_to)は、号を表す頂点を持つテキスト情報に基づいて構築する。

例：「平成12年3月4日」

- **事件番号**: 裁判所が取り扱う事件に付与されている番号。裁判所名・支部名・元号・事件記録符号・番号で構成される。

例：「東京地方裁判所／平成12年（ワ）第345678号」

- **掲載雑誌**: 判例が掲載されている雑誌(複数)。

例：「判例時報1234号567頁, 判例タイムズ1234号567頁」

- **参照法令**: 手作業で付与された、判決に関わる法令名と条項番号(複数)。

例：「特許法17の2-3, 特許法123-1」

1つの判例を1つの判例頂点とし、判例本文から抽出した上記の情報をメタデータとして追加する。判例から法令への引用関係を表す辺は、判例本文から抽出した参照法令名をもとに構築する。

2.4 法令

e-Gov法令検索^{*4}から、2024年9月時点で取得可能な法令XMLファイル約11,000件を対象とする。Louisらの研究[14]を参考に、法令は階層構造として表現する(図1右)。具体的には、以下の4種類の頂点を用いる：

- **法令頂点**: e-Gov法令検索より取得可能なXMLファイルの単位に相当する。法令名や制定文、公布日・改正日・施行日などのメタ情報を持つ。
- **条頂点**: 法令の条の単位に相当する頂点。子要素として項を持つ。条文見出しや条番号などのメタ情報を持つ。
- **項頂点**: 法令の項の単位に相当する頂点。子要素として号を持つ。項番号などのメタ情報を持つ。
- **号頂点**: 法令の号の単位に相当する頂点。テキスト情報をメタ情報として持つ。号以下の箇条書きの項目などのテキスト情報に関しては、号の単位に集約して1つのテキストとして扱う。

特筆すべき点として、法令は同一法令名であっても改正日や施行日異なる複数のバージョンを持つ。そのため、引用グラフ内ではこれらを区別して扱う。他の文献(書籍・ガイドライン・判例・法令)から法令への引用関係を構築する際には、引用元となる頂点を持つ日付情報(発行日・判決日・改正日等)に基づき、その時点における最新バージョンの法令へと対応付ける。

3 引用法令の抽出

引用法令の抽出については、比較的単純なルールベースであっても高い性能で抽出できることが先行研究により示されている[3, 20, 22]。本研究では、図2の引用法令抽出パイプラインにより、書籍・ガイドライン・判例・法令のテキスト中に出現する法令名を特定し、引用関係を表す辺を構築する。抽出対象の基本パターンを以下に示す：

- **法令番号**: 法令の種類・制定者・年別に付与される番号。
例：「平成十五年法律第五十七号」「昭和六十年郵政省令第二十五号」
- **法令名**: 法令の名称。
例：「個人情報の保護に関する法律」「電気通信事業法施行規則」

なお、いずれの表現においても、末尾に条・項・号の番号が付随する場合には、それらも含めて文字列を抽出する(例：「会社法第二十六条第一項」「会社法第三十二条第一項第一号」)。

これらの文字列パターンに加えて、次の2つの拡張処理を実装する：

- 略称表現への対応(3.1節)
- 基本パターンと条項番号の関係抽出(3.2節)

^{*4} <https://laws.e-gov.go.jp/>

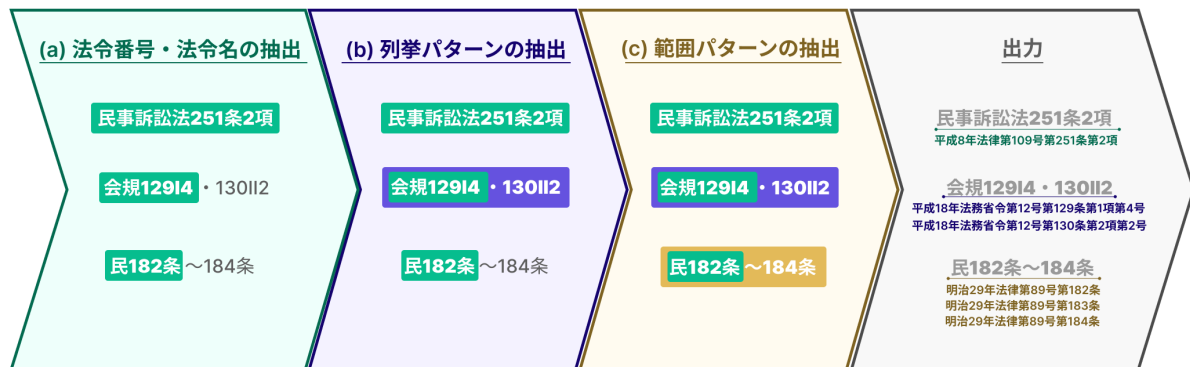


図2 引用法令抽出パイプラインにおける解析例. (a) 法令番号や法令名・法令名略称等の抽出 (b) 列挙パターンの抽出(3.2.1項) (c) 範囲パターンの抽出(3.2.2項)の順にルールを適用することで、引用法令文字列を抽出する。抽出した文字列に対して、法令番号や条項番号を補完する後処理を施すことにより、最終的な出力を得る。

3.1 略称表現への対応

3.1.1 書籍独自の略称

書籍においては、法令名や条項番号が独自の略称で表現されるケースが多い。例えば以下のようなケースが挙げられる。

- 会規124条3号:「会社法施行規則第124条第1項第3号」の省略形。
- 金商27条の2第1項3号:「金融商品取引法第27条の2第1項第3号」の省略形。
- 商登54IV:「商業登記法第54条第4項」の省略形。

これらの略称表現に対応するため、表1の形式の略称辞書を作成し、法令名の抽出ルールを拡張する。略称辞書は以下の情報を持つ:

- 略称: 法令名の略称。
- 正規の法令名: e-Gov法令検索から取得できるXMLファイル上に記載されている正式な法令名。
- コンテンツID: 書籍やガイドラインを一意に定める番号。文献ごとに略称方法は異なるため、略称の適用範囲を指定する。対象範囲を限定しない略称は「*」と記載する。

略称ルールによって抽出したパターンは、略称辞書を利用して正規の法令名に変換することで、法令名パターンとして扱う。

3.1.2 “法”による省略

書籍中で、引用法令が自明な場合は、単に「法」として参照するケースも存在する。例えば以下のようなケースが挙げられる。

- 法299条4項:「?法299条4項」を意味している。“?”に当てはまる法令名は書籍によって異なる。
- 法335III:「?法335条3項」を意味している。“?”に当てはまる法令名は書籍によって異なる。

正式法令名である“?”に当てはまる名称は、書籍タイトルや書籍概要などのメタ情報から具体的な法令名を抽出し、補完する。

3.2 関係抽出

3.2.1 条項の列挙

直前の法令名及び法令名略称を引き継ぐ形式で条項を列挙するケースを「条項の列挙」とよぶ。例えば以下のようなケースが挙げられる。

- 民法364条及び367条:「民法第364条」と「民法第367条」を意味している。
- 会更210条1項・214条:「会社更生法第210条第1項」と「会社更生法第214条」を意味している。

上記のケースには、列挙されている条項番号に対して、文字列先頭の法令名の補完処理により対応する。

3.2.2 条項の範囲指定

法令名や法令名略称と条項番号を範囲で記述するケースを「条項の範囲指定」とよぶ。例えば以下のようなケースが挙げられる。

- 民事再生法258条から260条:「民事再生法第258条」「民事再生法第259条」「民事再生法第260条」を意味している。
- 民138条～141条:「民法第138条」「民法第139条」「民法第140条」「民法第141条」を意味している。

上記のケースには、始点となる条項番号と終点となる条項番号を求め、その間の条項番号を列挙する後処理と法令名の補完処理により対応する。

4 引用判例の抽出

判例は一般に以下のような形式で引用することが多い:

- 東京地判平成18年3月31日判タ1209号60頁
- 東京地裁平成28年6月17日判決・判タ1436号201頁
- 最判昭61.2.18判時1185—151

これらの文字列は、以下の要素で構成されている [27]:

表1 略称辞書の例. 略称を適用する適用範囲を限定するため、書籍やガイドラインを一意に定めるコンテンツIDを付与した。ただし「*」と表記しているものは全コンテンツを対象とする。

法令略称	法令名	略称を適用するコンテンツID
LLP法	有限責任事業組合契約に関する法律	ABC001
金商	金融商品取引法	IJK002
外為法	外国為替及び外国貿易法	XYZ003
会規	会社法施行規則	*
商登則	商業登記規則	*
商登規	商業登記規則	*
商登	商業登記法	*
会更	会社更生法	*
民	民法	*

- 裁判所名: 裁判が行われた裁判所名(略記)
例: 「最高裁判所(最)」 「東京高等裁判所(東京高裁)」
- 裁判の種類: 判決(判), 命令(命), 決定(決)
- 裁判年月日: 裁判の年月日
例: 「平5.9.9」 「令和2年12月10日」
- 出典: 掲載されている判例集や雑誌(略称)とページ数
例: 「判例時報(判時)」 「判例タイムズ(判タ)」

引用判例抽出は、引用法令抽出パイプラインとは異なり、上記の要素で構成される文字列を抽出する正規表現のみで構成される。正規表現で文字列を抽出したのち、裁判所名や裁判の種類・出典などの略称を正式名称に戻すなどの後処理によって、引用判例を一意に特定する。

5 引用関係抽出処理の評価

5.1 評価データと評価指標

グラフ構築処理の評価のため、書籍・ガイドライン・法令からそれぞれ200チャンク（文書の一部）ずつ、計600チャンクを抽出し、それぞれのチャンク本文中に現れる法令名・判例名を手作業で付与した。これを評価データとして利用する。

評価指標として、Precision（適合率）、Recall（再現率）、F1 Scoreを採用し、提案したルールベース手法の性能を評価する。

引用法令抽出の評価では、条番号までを含むケースを対象とし、チャンクから適切な法令条文へ対応付けできた場合に正解とみなす。ただし、法令同士の引用関係において、同一法令内での引用は評価の対象外とする。

引用判例抽出の評価では、出典の先頭ページ番号までを対象とし、適切な判例へ対応付けできた場合に正解とみなす。

5.2 評価結果

各入力コンテンツごとの評価結果を表2に示す。判例から法令への引用関係は、判例本文に含まれる手作業で対応付けされた参照法令を用いたため、表2からは除外した。

表2の結果より、ルールベースの手法が法令や判例への引用関係の構築に高い効果を示していることが明らかとなった。

法令名抽出の結果に注目すると、Precisionが全体的に高く、すべての入力コンテンツでPrecisionが97%を超えている。特に法令

表2 ルールベースの法令名抽出・判例名抽出の評価結果. 判例から法令への引用関係の構築は判例本文に含まれる手作業で対応付けされた引用法令情報を利用したため、表からは除外した。また、法令から判例への引用は存在しないため未記載とした。

コンテンツ	引用法令抽出			引用判例抽出		
	Pre.	Rec.	F1	Pre.	Rec.	F1
書籍	97.75	83.39	90.00	85.71	80.77	83.17
ガイドライン	97.83	71.43	82.57	90.74	80.78	85.47
法令	98.80	92.13	95.35	-	-	-

から法令へのPrecisionは98.80%でF1 Scoreは95.35%と最も高い性能を示している。これは法令が形式的に記述されているため、ルールベースとの相性が良いことを示唆している。

書籍とガイドラインの法令名抽出の性能を比較すると、書籍の法令名抽出（F1 Score: 90.00%）がガイドラインの抽出（F1 Score: 82.57%）よりも高い性能を示している。この差は主にRecallの違いに起因しており、略称辞書により、書籍の幅広い表現を抽出できることが要因と考えられる。

判例名抽出においても、単純な正規表現で書籍・ガイドラインともにF1 Scoreが80%を上回る高い性能（83.17%と85.47%）を達成している。

全体を通して、PrecisionがRecallを上回っており、誤検出が少なく、抽出結果の信頼性が高いことを意味している。一方で、特にガイドラインにおける法令名抽出はRecall（71.43%）と、他のコンテンツにおける性能と比べて低く、改善の余地があることが分かる。

6 エラー分析

6.1 法令名抽出のエラー分析

書籍・ガイドラインにおける法令名抽出において、未検出だったものを目視で確認し、以下の3つのカテゴリに分類した：

1. テキストの前処理: 段組みや図表の読み取り処理、記号の正規化、文献のチャンク化などの前処理の際に法令名や条番号の情報が欠落したことによって、未検出となったもの。

表3 未検出エラーのカテゴリ別割合.

エラー	割合
テキストの前処理	5.7%
複雑な引用パターン	72.9%
未登録の法令名	21.4%

- 複雑な引用パターン: 条項の列挙や範囲指定の組み合わせや、「同法」をはじめとする直前の法令名参照など、複雑な引用パターンによって、未検出となったもの.
- 未登録の法令名: e-Gov法令検索では取得できない法令名や、略称辞書に登録されていない表現により、未検出となったものの.

未検出エラーのカテゴリ別割合(表3)を分析した結果、エラーの大部分(72.9%)は複雑な引用パターンによるものであった. このエラーの中でも、表4(b)のような「条項の列挙」と「条項の範囲指定」を交互に繰り返すケースや、(c)のように「同法」によって直前の法令名を継承するケースなどは、関係抽出の手法を高度化することで対応可能であり、性能向上の見込みがあることがわかった.

また、未登録の法令名によるエラーが約21%を占めていることから、略称辞書の拡張だけでも相当な改善が見込めることが示された. さらに、テキストの前処理に関するエラーは比較的小さいものの、前処理に関して改善の余地がまだあることも明らかになった.

6.2 判例名抽出のエラー分析

書籍・ガイドラインにおける判例名抽出のエラーを目視で確認したところ、主に未登録の出典略称や裁判所名略称に起因するものであった. これらは、引用判例抽出(4章)で述べた正規表現を拡張することで対応可能である. したがって、登録略称を定期的にメンテナンスすることで、高い精度で判例名抽出が可能であることが明らかになった.

これらの分析結果から、法令名抽出と判例名抽出の両方において、辞書やルールの拡張という比較的単純な方法でも、さらなる精度向上が期待できることが示唆された.

7 引用法令抽出ルールごとの効果検証

3章で提案した引用法令抽出ルールの有効性を検証するため、Ablation Studyを実施した. 具体的には、以下の4つのルールをパイプラインから個別に除去し、それぞれのルールの効果を評価した:

- 書籍独自の略称(3.1.1項)
- “法”による省略(3.1.2項)
- 条項の列挙(3.2.1項)
- 条項の範囲指定(3.2.2項)

結果を表5に示す. なお、すべてのルールを適用した場合(表中の「除外したルールなし」)の結果は、表2の書籍における引用法令抽出の結果と一致する.

分析結果から「書籍独自の略称」および「“法”による省略」の

ルールが特に重要であることが分かる. これらのルールを除外すると、Recallが27~30ポイント程度低下した.

また「“法”による省略」のルールを除外した場合、Precisionがわずかに向上している. これは、このルールが一部の誤検出を引き起こしている可能性を示唆している. その一方で、「“法”による省略」のルールを除外した場合にRecallが大幅に低下していることから、このルールがRecallの向上に大きく貢献していることが分かる. すべてのルールを適用した場合と比較して、各ルールを除外した場合のPrecisionの変動は比較的小さい. これは、提案したルールが高いPrecisionを維持しつつ、Recallを向上させる効果があることを示している.

8 今後の展望

本研究で構築した引用グラフは2024年9月時点のものであり、今後も継続的に発展させていく予定である. また、本研究で構築した引用グラフを法律分野の様々なタスクへ応用することを計画している. 例えば、法律文書分類[5, 17, 19]、類似判例検索[4, 21]、法律関連の質問応答システム[7, 15]、法的判断予測[2, 23]などのタスクで活用する予定である.

これらの取り組みを通じて、本研究で構築した引用グラフの有用性をさらに高め、法律分野の研究と実務の両面で、より効果的かつ効率的な情報収集・分析が可能となることを目指す.

参考文献

- Jiaxi Cui, Munan Ning, Zongjian Li, Bohua Chen, Yang Yan, Hao Li, Bin Ling, Yonghong Tian, and Li Yuan. Chatlaw: A multi-agent collaborative legal assistant with knowledge graph enhanced mixture-of-experts large language model. *arXiv [cs.CL]*, June 2023.
- Junyun Cui, Xiaoyu Shen, and Shaochun Wen. A survey on legal judgment prediction: Datasets, metrics, models and challenges. *IEEE Access*, 11:102050–102071, 2023.
- Émile de Maat, Radboud Winkels, and Tom van Engers. Automated detection of reference structures in law. In *Proceedings of the 2006 conference on Legal Knowledge and Information Systems: JURIX 2006: The Nineteenth Annual Conference*, pages 41–50, NLD, June 2006. IOS Press.
- Jaspreet Singh Dhani, Ruchika Bhatt, Balaji Ganesan, Parikshet Sirohi, and Vasudha Bhatnagar. Similar cases recommendation using legal knowledge graphs. In *SAIL' 23: 3rd Symposium on Artificial Intelligence and Law*, February 2023.
- Tuggener Don, Pius von Däniken, Peetz Thomas, and Cieliebak Mark. LEDGAR: A large-scale multi-label corpus for text classification of legal provisions in contracts. In Nicoletta Calzolari, Frédéric Bécet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Twelfth Language Resources and Evaluation Conference*, page 1235–1241. European Language Resources Association, May 2020.
- Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, and Jonathan Larson. From local to global: A graph RAG approach to query-focused summarization. *arXiv [cs.CL]*.
- Yi Feng, Chuanyi Li, and Vincent Ng. Legal case retrieval: A survey of the state of the art. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6472–6485, Stroudsburg, PA, USA, August 2024. Association for Computational Linguistics.
- Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. Retrieval-augmented generation for large language models: A survey. *arXiv*

表4 引用法令抽出エラーの具体例。抽出結果のうち未検出のものを目視で確認したのち、3つのカテゴリに分類した。具体例テキスト中の下線部が正解の引用法令を表す文字列である。また、テキストの前処理によって注釈が本文と混ざってしまった部分には打ち消し線を付与した。

	エラー	具体例	予測結果	正解
(a)	テキストの前処理	...労働条件の維持・改善その他経済的地位の向上を図ることにあり(労働2監査役資格および選任→終任2)その活動自体は...	-	労働組合法第22条
(b)	複雑な引用パターン	破産法265条,266条,268条~272条,274条これらの罪により刑に処せられ,その執行を終わり,またはその執行を受けることがなくなった日から2年を...	破産法265条	破産法266条・破産法268条・破産法269条・破産法270条・破産法271条・破産法272条・破産法274条
(c)	複雑な引用パターン	...株式併合の差止め(会社法182条の3),株式等売渡請求に係る売渡株式等の取得の差止め(同法179条の7)...	会社法182条の3	会社法182条の3・会社法179条の7
(d)	未登録の法令名	...に規定する参考書類には,上記と同様の記載が必要となる(勧誘府令4IV)。	-	上場株式の議決権の代理行使の勧誘に関する内閣府令4条4項

表5 各ルールを除外した場合の性能比較。

除外したルール	法令名抽出		
	Pre.	Rec.	F1
なし	97.75	83.39	90.00
書籍独自の略称	97.27	56.87	71.78
“法”による省略	98.81	53.04	69.03
条項の列挙	97.72	68.69	80.67
条項の範囲指定	97.64	79.55	87.67

[cs.CL], December 2023.

- [9] Xiaoxin He, Yijun Tian, Yifei Sun, Nitesh V Chawla, Thomas Laurent, Yann LeCun, Xavier Bresson, and Bryan Hooi. G-retriever: Retrieval-augmented generation for textual graph understanding and question answering. *arXiv [cs.LG]*, February 2024.
- [10] Yuntong Hu, Zhihan Lei, Zheng Zhang, Bo Pan, Chen Ling, and Liang Zhao. GRAG: Graph retrieval-augmented generation. *arXiv [cs.LG]*, May 2024.
- [11] Bowen Jin, Chulin Xie, Jiawei Zhang, Kashob Kumar Roy, Yu Zhang, Zheng Li, Ruirui Li, Xianfeng Tang, Suhang Wang, Yu Meng, and Jiawei Han. Graph chain-of-thought: Augmenting large language models by reasoning on graphs. In *Martins Srikumar Lun-Wei Ku, editor, In Findings of the Association for Computational Linguistics: ACL 2024*, page 163–184. Association for Computational Linguistics, August 2024.
- [12] Jaewoong Kim and Moohong Min. From RAG to QA-RAG: Integrating generative AI for pharmaceutical regulatory compliance process. *arXiv [cs.CL]*, January 2024.
- [13] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Kuttler, Mike Lewis, Wen-Tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, May 2020.
- [14] Antoine Louis, Gijs van Dijck, and Gerasimos Spanakis. Finding the law: Enhancing statutory article retrieval via graph neural networks. In *Andreas Vlachos and Isabelle Augenstein, editors, Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, page 2761–2776, Dubrovnik, Croatia. Association for Computational Linguistics.
- [15] Antoine Louis, Gijs van Dijck, and Gerasimos Spanakis. Interpretable long-form legal question answering with retrieval-augmented large language models. *arXiv [cs.CL]*, September 2023.
- [16] Costas Mavromatis and George Karypis. GNN-RAG: Graph neural retrieval for large language model reasoning. *arXiv [cs.CL]*, May 2024.
- [17] Limsopatham Nut. Effectively leveraging BERT for legal document classification. In *Aletras Nikolaos, Androustopoulos Ion, Barrett Leslie, Goanta Catalina, and Daniel Preotiuc-Pietro, editors, Proceedings of the Natural Language Processing Workshop 2021*, page 210–216. Association for Computational Linguistics, 2021.
- [18] Boci Peng, Yun Zhu, Yongchao Liu, Xiaohe Bo, Haizhou Shi, Chuntao Hong, Yan Zhang, and Siliang Tang. Graph retrieval-augmented generation: A survey. *arXiv [cs.AI]*, August 2024.
- [19] Wang Qiqi, Zhao Kaiqi, Amor Robert, Liu Benjamin, and Wang Ruofan. D2GCLF: Document-to-graph classifier for legal document classification. In *Carpuat Marine, Marie-Catherine de Marneffe, Vladimir Ivan, and Ruiz Meza, editors, Findings of the Association for Computational Linguistics: NAACL 2022*, page 2208–2221. Association for Computational Linguistics, July 2022.
- [20] Ali Sadeghian, Lakshman Sundaram, Daisy Zhe Wang, William F Hamilton, Karl Branting, and Craig Pfeifer. Automatic semantic edge labeling over legal citation graphs. *Artif. Intell. Law*, 26(2):127–144, June 2018.
- [21] Yanran Tang, Ruihong Qiu, Hongzhi Yin, Xue Li, and Zi Huang. CaseLink: Inductive graph learning for legal case retrieval. In *The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, March 2024.
- [22] Oanh Thi Tran, Minh Le Nguyen, and Akira Shimazu. Reference resolution in legal texts. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, New York, NY, USA, June 2013. ACM.
- [23] Hiroaki Yamada, Takenobu Tokunaga, Ryutaro Ohara, Akira Tokutsu, Keisuke Takeshita, and Mihoko Sumida. Japanese tort-case dataset for rationale-supported legal judgment prediction. *Artif. Intell. Law*, May 2024.
- [24] 久本空海, 城戸祐亮, 津金澤佳亨, and 八木田樹. 民事判決のオープンデータ化へ向けた機械処理による判例仮名化の検証. In *言語処理学会第28回年次大会発表論文集*, pages 1406–1410, March 2022.
- [25] 前田郁勝, 外山勝彦, and 小川泰弘. 民事第一審判決書のXMLデータ化. In *言語処理学会第30回年次大会発表論文集*, pages 1324–1328, March 2024.
- [26] 山内匠. 法令データの現状と法令分野へのデジタル技術適用の展望. In *言語処理学会第30回年次大会発表論文集*, pages 1051–1055, March 2024.
- [27] 法律編集者懇話会. 法律文献等の出典の表示方法. 2014.