

自然言語処理に適した法令表記に関する調査研究 報告書

調達番号：令和6年度 0049-0226

アイビー・システム株式会社

令和7年3月31日

目次

1. 本調査研究の目的	3
2. 調査範囲	3
3. 調査概要	4
3. 1. 法令中の併合的接続詞・選択的接続詞を用いた表記に関する調査	7
3. 1. 1. 現状分析と調査方針	7
3. 1. 2. 調査方法	11
3. 2. 法令中の定義語・略称表記に関する調査	15
3. 2. 1. 現状分析と調査方針	15
3. 2. 2. 調査方法	16
4. 法令中の併合的接続詞・選択的接続詞の表記に関する調査について整理分析	16
4. 1. 「条文生成 UI」で新規に成文化する機能の実現可能性について	16
4. 1. 1. 読点による接続の課題	17
4. 1. 2. 入力操作の課題	18
4. 2. 「条文生成 UI」で既存条文を扱う実現可能性について	20
4. 2. 1. 各接続詞が接続対象とする要素及び範囲の特定について	20
4. 2. 2. 各接続詞が条文中で属する入れ子階層の特定	21
4. 3. データ構造に起因する課題の整理	21
4. 3. 1. 課題の種類一覧	21
4. 3. 2. 課題の種類に関する事例	22
4. 4. 法令の表記ルールに起因する課題の整理	22
4. 4. 1. 課題の種類	22
4. 4. 2. 課題の種類に関する事例	23
4. 5. 内容解釈に起因する課題の整理	24
4. 5. 1. 課題の種類一覧	24
4. 5. 2. 課題の種類に関する事例	24
4. 6. 例外的な表現について	25
4. 7. 調査結果に基づく考察	26
5. 法令中の定義語・略称表記に関する調査についての整理分析	27
5. 1. 「以下リスト UI」の実現可能性について	27
ステップ1：自然言語処理による定義語・略称の識別が可能な表記の種類（パターン）の 特定	28
ステップ2：パターンごとに法令をサンプリング	29
ステップ3：「以下リスト UI」生成用データの作成と実装	30
5. 2. 「以下リスト UI」の機能実現のために自然言語処理が活用できる範囲について	32

5. 2. 1. 定義条文におけるパターン分類.....	32
5. 2. 2. 略語表記におけるパターン分類.....	34
5. 3. 法令の表記ルールに起因する課題の整理.....	35
5. 3. 1. 課題の類型一覧.....	35
5. 3. 2. 課題の類型に関する事例.....	36
5. 4. 内容解釈に起因する課題の整理.....	39
5. 4. 1. 課題の類型一覧.....	39
5. 4. 2. 課題の類型に関する事例.....	39
5. 5. 調査結果に基づく考察.....	40
6. 本調査研究のまとめ及び今後の提案.....	41
6. 1. 法令中の併合的・選択的接続詞の表記への自然言語処理のアプローチについて.....	41
6. 2. 定義語・略称に対するリンク情報付与への自然言語処理のアプローチについて.....	42
6. 3. まとめ.....	43

1. 本調査研究の目的

本レポートは、令和6年度0049-0226「自然言語処理に適した法令表記に関する調査研究」仕様書に基づく調査研究に係る最終報告書である。

総務省は、「法案誤り等再発防止プロジェクトチーム取りまとめ」（令和3年6月29日）及び「デジタル社会の実現に向けた重点計画」（令和6年6月21日閣議決定）等を踏まえ、デジタル庁及び関係機関と連携しつつ、デジタル技術・ICTを積極的に活用する形での法制執務全体の業務フローの在り方に関する検証に取り組んでいるところである。

上記のような状況から、業務効率化・誤り防止の観点からニーズの高い課題として、「①併合的接続詞・選択的接続詞（及び・並びに、又は・若しくは）」、「②定義語・略称表記」が挙げられる。これらの課題に起因する法制執務に当たる職員の負担は大きい。例えば「①併合的接続詞・選択的接続詞（及び・並びに、又は・若しくは）」については接続詞の使い方が複雑な条文においてその用法が正しく用いられているかを目視チェックで確認するなどの業務負担が生じており、「②定義語・略称表記」についても定義語・略称を定めた条文とそれらを用いている条文の対応関係が過不足なく規定できているか確認するために目視チェックで規定の対応関係を把握しリスト管理している（いわゆる「以下リスト」）など、大きな作業負担がかかっているのが実情である。

本調査研究は、これらの課題を踏まえ、「自然言語処理」の活用による法制執務の業務効率化や誤りの発生リスクの低減への影響と、その前提となる法令表記の機械可読化の見直し的一端として、自然言語処理に適した法令表記ルールの在り方についての調査・実証を行うものである。

2. 調査範囲

本調査研究では「自然言語処理に適した法令表記ルール」を調査する対象として、本調査研究のキックオフ実施時点（令和6年11月15日）でe-Gov法令検索からダウンロードできる全法令XMLデータ（11,508法令）のうち

- ・未施行の法令を除外（11,508法令→9,229法令）
- ・文語体で書かれた法令を除外（9,229法令→8,956法令）

8,956 法令を分析対象とした。また、対象となった法令の内訳は以下のとおりである。

- ・憲法 1 件
- ・勅令 14 件
- ・法律 2,020 件
- ・政令 2,359 件
- ・府省令 4,146 件
- ・規則 416 件

これらの分析対象に対し、前述した以下の表現に限定し調査を行った。

- ① 併合的接続詞・選択的接続詞（及び・並びに、又は・若しくは）
- ② 定義語・略称

また、法令表記に関する基本的な解釈については「改訂 法令解釈の基礎」（長谷川彰一著 株式会社ぎょうせい）、「法制執務詳解 新版Ⅲ」（石毛正純著 株式会社ぎょうせい）及び「ワークブック法制執務 第2版」（法制執務研究会編 株式会社ぎょうせい）を参照した。

3. 調査概要

日本語は、以下の特徴から、他の言語と比較して、その文意や文脈を機械判読することが困難な言語である。

- ・語の間にスペースを持たないため、語の境界を判別することが困難
- ・語順が固定されておらず、語を入れ替えても意味が通じる
- ・係る語と係られる語の間に無限に修飾語を入れることができる

このため法令条文を機械処理するためには、条文のテキストを形態素解析¹し、意味を持つ最小単位の語に分解するとともに、その語がどのような品詞であるのか、どの語に対して係り受け関係を持つのかという情報を付与する必要がある。このような「語の分割」から「品詞情報の付与」、「語の係り受け解析」の一連の処理を「自然言語処理」と呼ぶ。

¹ 文章を、意味を持つ最小の単位（形態素）に分解し、品詞や活用などの情報を付与する処理

本調査研究では、調査・分析を進める中で法令条文を自然言語処理し、品詞情報・係り受け情報を得る必要があったため、自然言語処理用オープンソース²ソフトウェア・ライブラリ³spaCy⁴、spaCyをベースとした日本語向け自然言語処理オープンソースソフトウェア・ライブラリ GiNZA等以下のソフトウェアを使用し、本来テキストデータ等に出力される結果をwebブラウザによって視覚的に品詞情報・係り受け情報を把握できる「表記パターン解析ツール」（「別添資料 主要プログラムコード」中の「表記パターン解析ツール」参照）を開発し利用した。

anaconda⁵: 24.11.3

python⁶: 3.12.7.final.0

spaCy: 3.7.2

GiNZA⁷(ja-ginza-electra): 5.2.0

streamlit⁸: 1.40.1

開発したツールは、入力された文章に対し「分かち書き（語の分解）」、「品詞情報の付与」、「係り受け解析」の各処理がなされ、法令条文内でそれらの結果を利用した「併合的接続詞・選択的接続詞が用いられる表現」や「定義語・略称に関連する表現」のパターン抽出プログラムや、係り受け情報から文章内容を把握するためのプログラムのロジックに関する検討等に活用した。

² オープンソースとは、不特定多数の人が活用・改良することを目的としてソフトウェアのソースコードを公開し、誰でも自由に利用や改変、再配布できることを指す

³ 特定の処理を行うためのプログラムを、他のソフトウェアから容易に利用できるように部品化したもの

⁴ 字句解析処理のソフトウェア・ライブラリ。<https://spacy.io/>

⁵ 科学計算のソフトウェア・ライブラリ。<https://www.anaconda.com/>

⁶ プログラミング言語のひとつ、及びその実行環境。<https://www.python.org/>

⁷ 日本語向け自然言語処理のソフトウェア・ライブラリ。<https://megagonlabs.github.io/ginza/>

⁸ 機械学習向けデータ視覚化のソフトウェア・ライブラリ。<https://streamlit.io/>

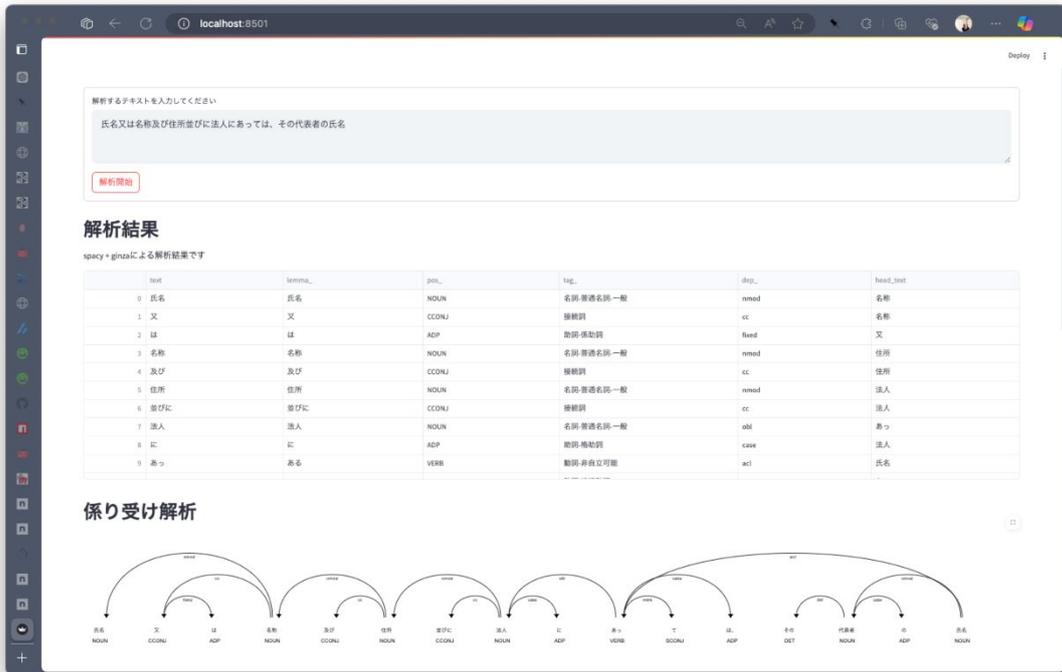


図1 表記パターン解析ツール画面

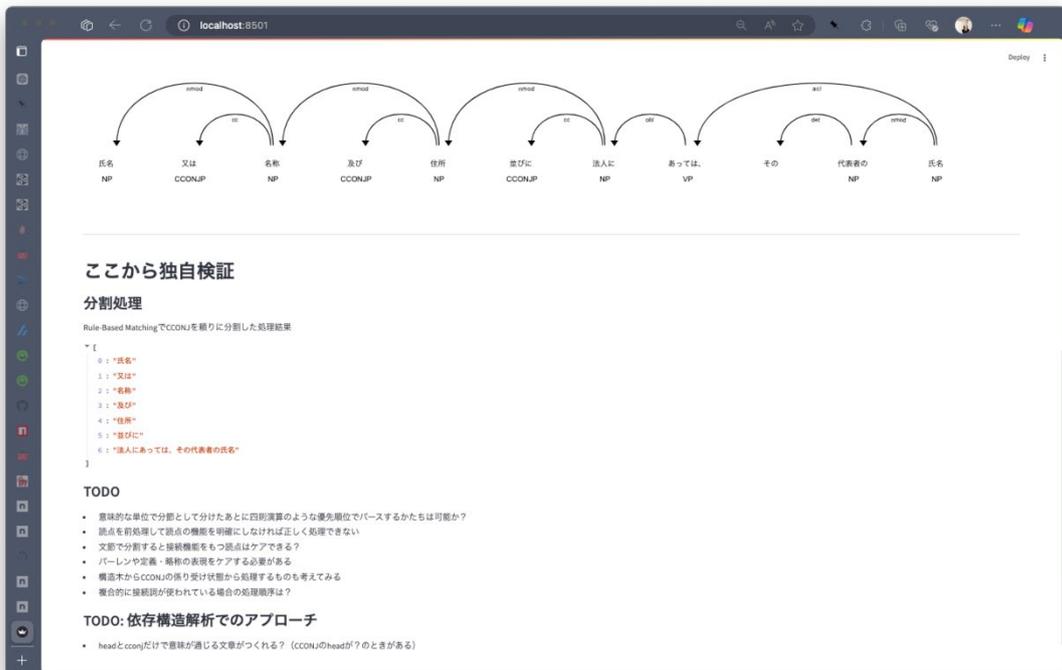


図2 表記パターン解析ツール画面

また、法令条文の「併合的接続詞・選択的接続詞を用いた表記」、「定義語・略称表記」はそれぞれの特徴があり、その特徴に基づいた表記ルールの分析調査を行う必要がある。それらの特徴及びその特徴を考慮した調査方法については次項にて示す。

3. 1. 法令中の併合的接続詞・選択的接続詞を用いた表記に関する調査

3. 1. 1. 現状分析と調査方針

法令条文中の併合的接続詞として用いられる「及び」、「並びに」、選択的接続詞として用いられる「又は」、「若しくは」の使用については、以下のように基本的なルールが存在する。

併合的接続詞「及び」、「並びに」

- 「及び」は常に最小単位の接続となる
- 並列をさらに細分化して結び付ける必要がある場合には「並びに」が使用される
- 3つ以上の名詞が接続される場合には「、」（読点）を使う（e.g. 「A、B、C及びD」）
- 接続されるものが動詞・形容詞・副詞の場合はその語が2つでも「、」（読点）を使う（e.g. 「輸出し、及び輸入する」）

選択的接続詞「又は」、「若しくは」

- 「又は」は常に一番大きな接続となる
- 選択をさらに細分化して結び付ける必要がある場合には「若しくは」が使用される
- 3つ以上の名詞が接続される場合には「、」（読点）を使う（e.g. 「A、B、C又はD」）
- 接続されるものが動詞・形容詞・副詞の場合はその語が2つでも「、」（読点）を使う（e.g. 「忌避し、又は同項の規定による」）

この基本的なルールに加え、実際的にはこれら併合的接続詞と選択的接続詞が複合的に使われる場合や、並列・選択を多段的に結び付ける場合があるため、どの語句とどの語句を論理積又は論理和として接続しているのかが分かりづらく、人間可読性及び機械可読性の観点から課題がある。

このため本調査研究では、併合的接続詞・選択的接続詞が用いられる条文について、自然言語処理による解析情報を含んだ機械可読性の高い構造化データを用いることにより、人間可読性の高い「意味を成す語のまとまりを文章以外の表現で視覚化する方法」（以下「文章のオブ

ジェクト表現」という。)について提案した。「文章のオブジェクト表現」の提案については、「新規条文を成文化する場合」と「既存の法令条文に適用する場合」とに分け、これらの実現可能性について調査・実証し、課題の類型化・整理分析を行った。

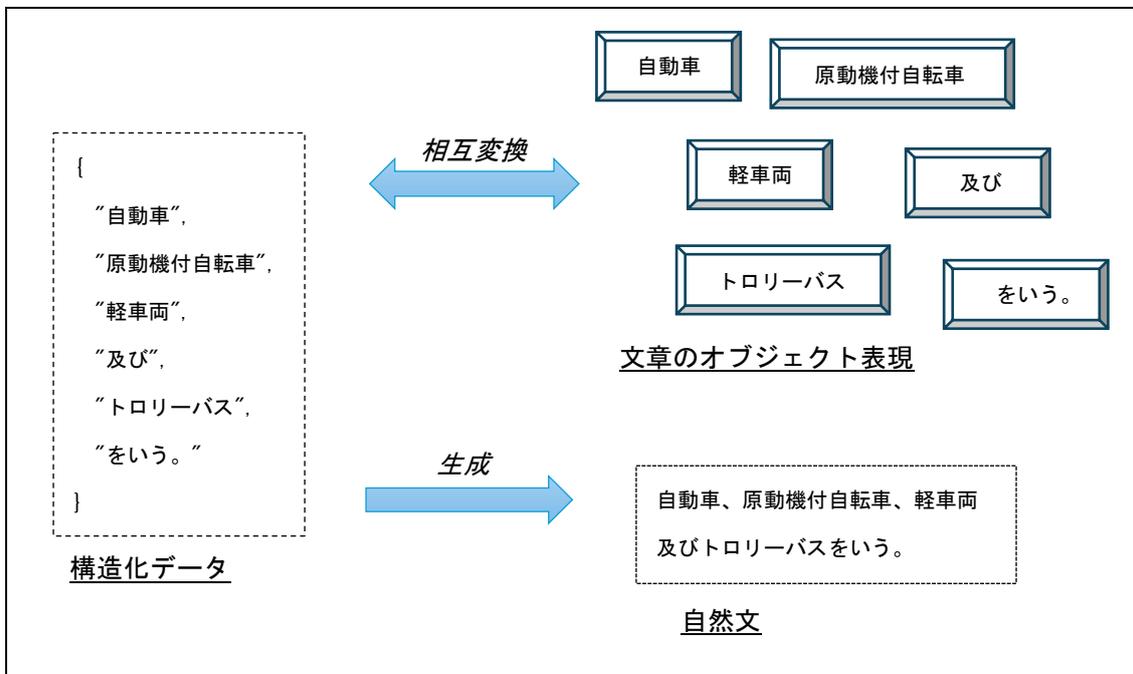


図3 構造化データと「文章のオブジェクト表現」の活用イメージ

以下は Web ブラウザ環境を前提とした「文章のオブジェクト表現」の例である。

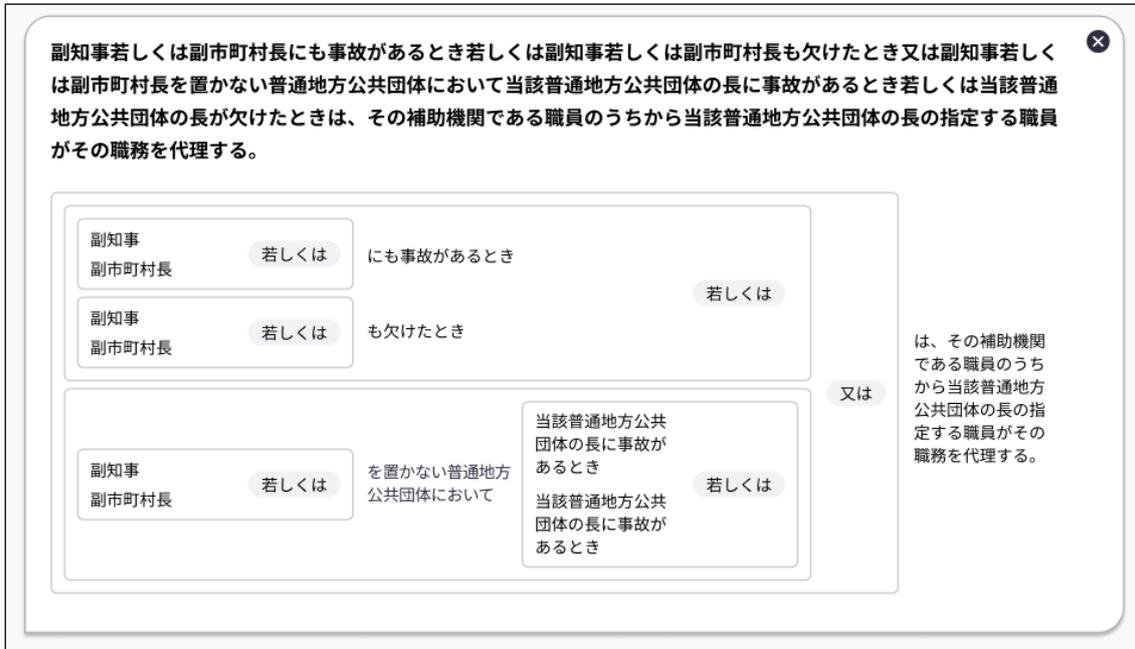


図 4 「文章のオブジェクト表現」の例

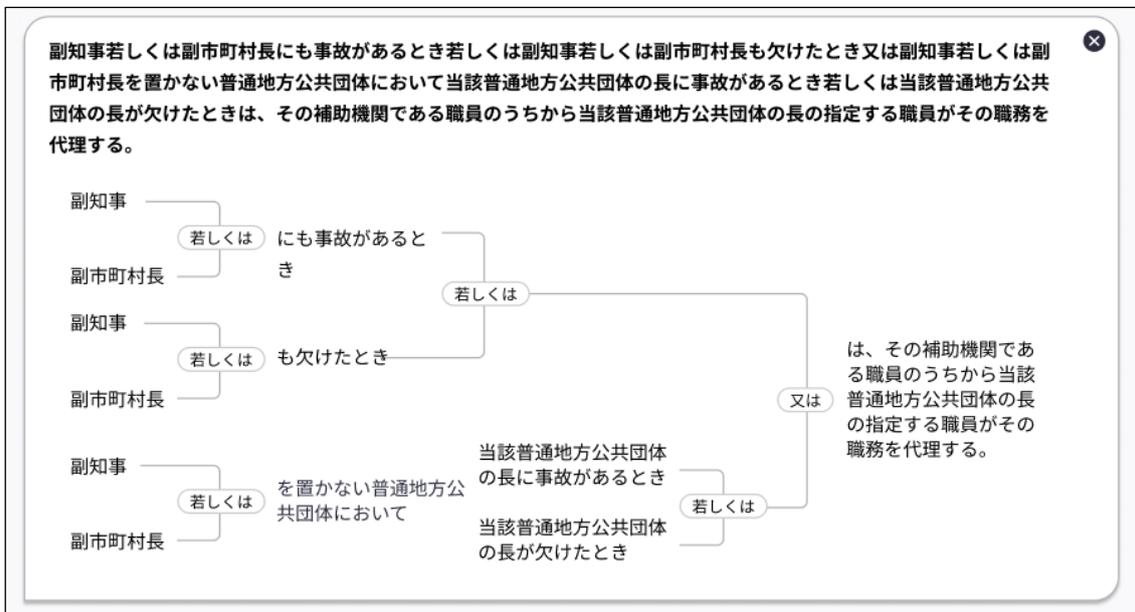


図 5 「文章のオブジェクト表現」の例

図6 「文章のオブジェクト表現」の例（入力フォーム）

「新規条文を成文化する場合」と「既存の法令条文に適用する場合」に分けて調査・実証及び分析を行った理由であるが、まず、新規条文を成文化する場合、複雑で長大な文章でない限りは、フォームのようなインターフェースを利用したユーザー入力による「構造化データ」から「自然文」を構築する問題と考えることができる。この方向性により「接続詞」をヒントとした文字列処理的な観点から調査・実証をはじめることができ、その方針での処理できる限界とそれをカバーするための自然言語処理の導入についての検討ができるのではと想定した。

一方、既存の法令条文については、新規の成文化とは逆に「自然文」から「構造化データ」を構築する問題となる。この場合、自然文である法令条文を自然言語処理することにより「オブジェクト表現」のための「構造化データ」を作成しなければならないため、事前に併合的接続詞・選択的接続詞が用いられる条文を類型化（パターン化）し、自然言語処理が困難なものについての分析が必須になり、新規の成文化に比べ難易度が高まり、調査・検証を分けて行う必要があると想定した。

3. 1. 2. 調査方法

はじめに「新規条文を成文化する場合」であるが、デジタル庁・総務省における既往の検討の中で議論されていた新規条文を「エディタの中で読み出す機能」は、Figma⁹上でプロトタイプとして動作するものであったが、本調査研究では、これを Web ブラウザベースで動作する形へ開発・実装していく段階を想定し、技術的な観点からだけでなく、法令表記の観点からどのような検討課題が生じるかをまとめ、「新規条文を成文化する場合」に用いる UI¹⁰（以下「条文生成 UI」という。）がもたらす「併合的接続詞と選択的接続詞を用いた表現」の「文章のオブジェクト表現」による可読性の向上の可能性について整理・分析する方法をとることとした。

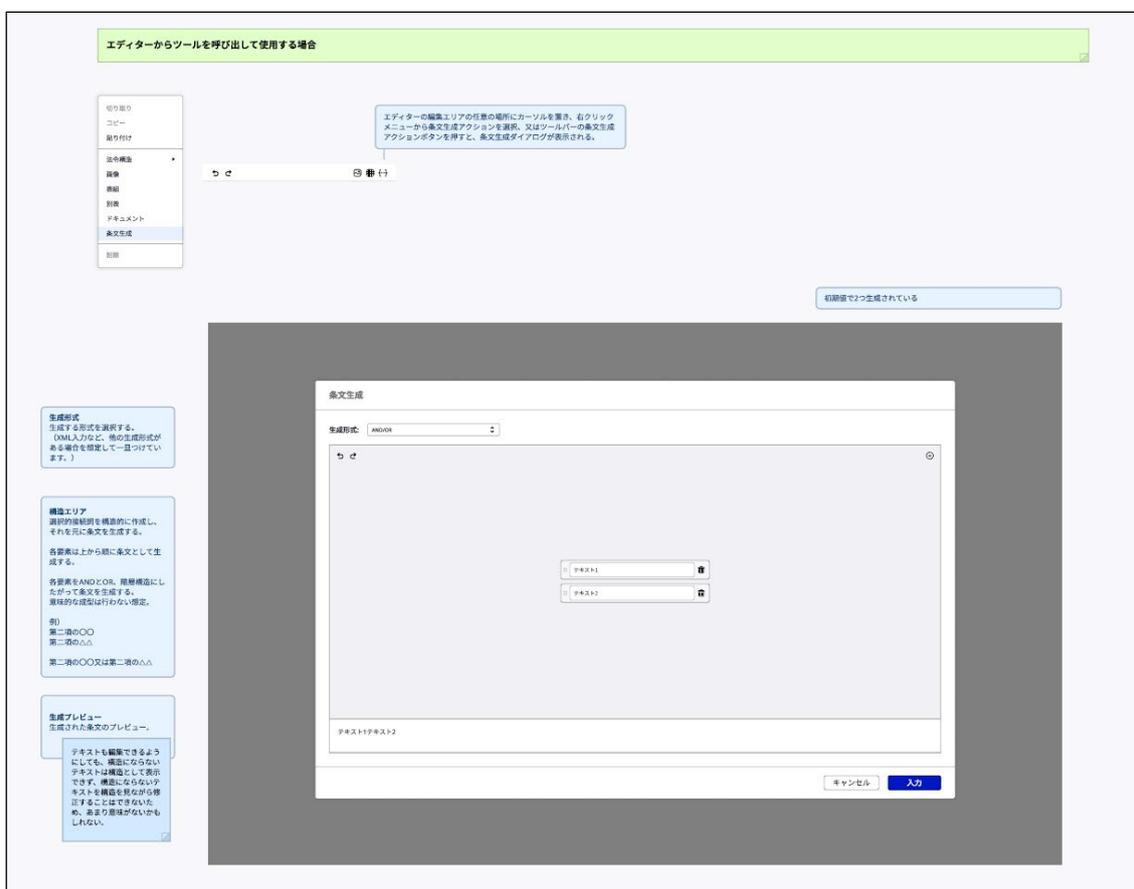


図7 FigmaによるプロトタイプのUI検証

⁹ 各種部品を配置するなどして、画面のデザインを検討できる Web ツール。

<https://www.figma.com/ja-jp/>

¹⁰ ユーザー・インターフェースの略。ここでは、成文化作業で用いる操作画面を指す

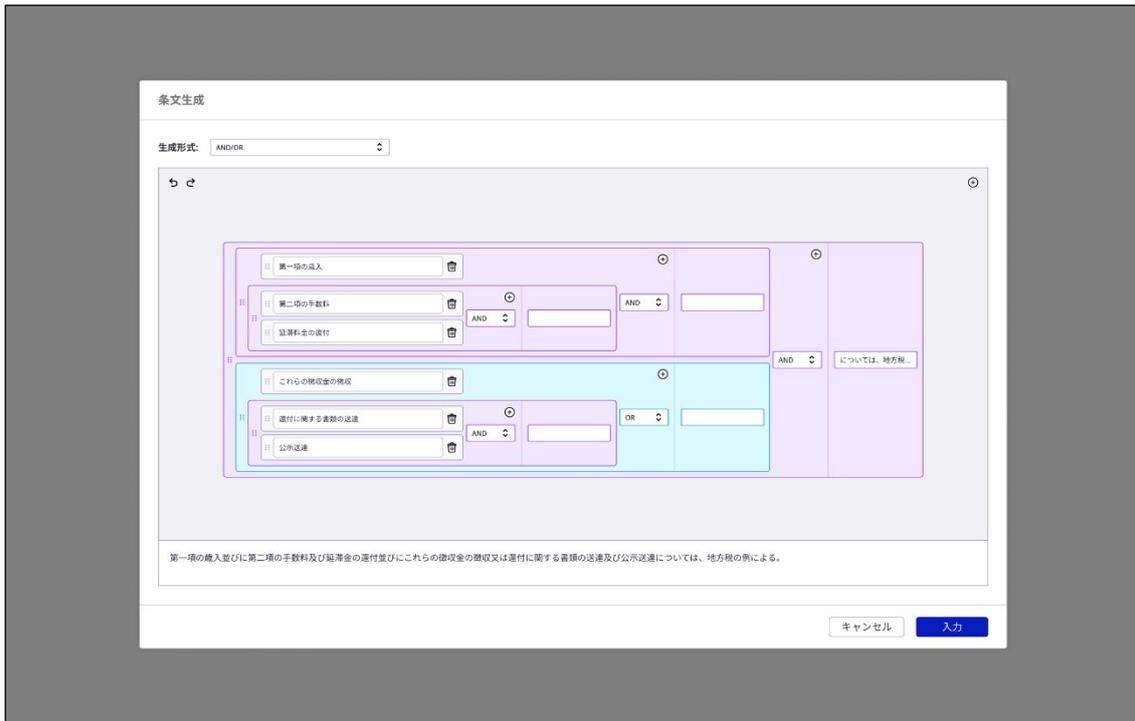


図8 FigmaによるプロトタイプのUI検証



図9 FigmaによるプロトタイプのUI検証

次に「既存の法令条文に適用する場合」については、「文章のオブジェクト表現」を実現するため自然言語処理を実施し、品詞や係り受けに関する情報を付与する必要があった。このため自然言語処理を困難とするような表記パターンがどの程度存在するかを把握・整理する必要があり、まず、「併合的接続詞と選択的接続詞を用いた表現」を機械的に抽出した。

本調査研究においては分析対象となる法令条文のうち本則に絞って抽出を行った。本則に絞った理由としては、附則・別表においては、接続詞を含む文章の一部が「準用」や「読み替え」などのより複雑な法令表記の中で用いられるようなことが多いこと、文字列だけでなく図表による併合的接続詞と選択的接続詞を用いるようなものが多いなど、本則に比べ特殊な表記が多かったためである。抽出は法令 XML 中の Sentence タグ内¹¹で併合的接続詞として用いられる「及び」、「並びに」、選択的接続詞として用いられる「又は」、「若しくは」をカウントした。以下が抽出した結果である。

- 調査対象法令の全センテンス数： 3,439,036 センテンス
- 併合的・選択的接続詞を含むセンテンス数： 975,859 センテンス（全センテンスのおよそ 28.4%）
- 併合的・選択的接続詞を含む「本則」のセンテンス数： 463,397 センテンス（全センテンスのおよそ 13.5%）
- 併合的接続詞「及び」、「並びに」のみ使われている本則センテンス数： 184,965 センテンス（全センテンスのおよそ 5.4%）
- 選択的接続詞「又は」、「若しくは」のみ使われている本則センテンス数： 202,652 センテンス（全センテンスのおよそ 5.9%）

また、法令種別毎の割合は以下のとおりである。

憲法

- 「憲法」の全センテンス数: 219 センテンス
- 併合的・選択的接続詞を含む「本則」のセンテンス数: 76 センテンス（「憲法」の全センテンスのおよそ 34.7%）
- 併合的接続詞「及び」「並びに」のみ使われている本則センテンス数: 43 センテンス（「憲法」の全センテンスのおよそ 19.6%）
- 選択的接続詞「又は」「若しくは」のみ使われている本則センテンス数: 31 センテンス（「憲法」の全センテンスのおよそ 14.2%）

勅令

- 「勅令」の全センテンス数: 4,168 センテンス
- 併合的・選択的接続詞を含む「本則」のセンテンス数: 1,086 センテンス（「勅令」の全センテンスのおよそ 26.1%）

¹¹ 法令標準 XML スキーマ <https://laws.e-gov.go.jp/help/#law-xml-schema>

- 併合的接続詞「及び」「並びに」のみ使われている本則センテンス数: 387 センテンス（「勅令」の全センテンスのおよそ 9.3%）
- 選択的接続詞「又は」「若しくは」のみ使われている本則センテンス数: 492 センテンス（「勅令」の全センテンスのおよそ 11.8%）

法律

- 「法律」の全センテンス数: 938,874 センテンス
- 併合的・選択的接続詞を含む「本則」のセンテンス数: 175,003 センテンス（「法律」の全センテンスのおよそ 18.6%）
- 併合的接続詞「及び」「並びに」のみ使われている本則センテンス数: 56,625 センテンス（「法律」の全センテンスのおよそ 6.0%）
- 選択的接続詞「又は」「若しくは」のみ使われている本則センテンス数: 90,343 センテンス（「法律」の全センテンスのおよそ 9.6%）

政令

- 「政令」の全センテンス数: 631,645 センテンス
- 併合的・選択的接続詞を含む「本則」のセンテンス数: 92,796 センテンス（「政令」の全センテンスのおよそ 14.7%）
- 併合的接続詞「及び」「並びに」のみ使われている本則センテンス数: 39,369 センテンス（「政令」の全センテンスのおよそ 6.2%）
- 選択的接続詞「又は」「若しくは」のみ使われている本則センテンス数: 37,862 センテンス（「政令」の全センテンスのおよそ 6.0%）

府省令

- 「府省令」の全センテンス数: 1,712,481 センテンス
- 併合的・選択的接続詞を含む「本則」のセンテンス数: 180,347 センテンス（「府省令」の全センテンスのおよそ 10.5%）
- 併合的接続詞「及び」「並びに」のみ使われている本則センテンス数: 82,641 センテンス（「府省令」の全センテンスのおよそ 4.8%）
- 選択的接続詞「又は」「若しくは」のみ使われている本則センテンス数: 68,325 センテンス（「府省令」の全センテンスのおよそ 4.0%）

規則

- 「規則」の全センテンス数: 151,649 センテンス
- 併合的・選択的接続詞を含む「本則」のセンテンス数: 14,089 センテンス（「規則」の全センテンスのおよそ 9.3%）

- 併合的接続詞「及び」「並びに」のみ使われている本則センテンス数: 5,900 センテンス（「規則」の全センテンスのおよそ 3.9%）
- 選択的接続詞「又は」「若しくは」のみ使われている本則センテンス数: 5,599 センテンス（「規則」の全センテンスのおよそ 3.7%）

このように抽出した条文から、自然言語処理を行う場合に課題となる可能性があるものについての調査を進め、「自然文」から「構造化データ」への展開について、特に「条文生成 UI」で既存条文を表現する場合についての考察を行い、併せて法令の表記ルールに関する課題について整理分析を行うこととした。

3. 2. 法令中の定義語・略称表記に関する調査

3. 2. 1. 現状分析と調査方針

法令条文において、定義語・略称を用いる際には『以下「○○』という。』等のように表記される。法令の立案段階においては、定義語・略称を含む法令を立案・改正する場合、条文内の定義語・略称の表記の適正化を図る観点から、定義語・略称の定義を行った条文と定義された定義語・略称が使用されている条文を手作業で検索し目視確認をするなど、業務負担が高い方法で把握・管理している点で課題がある。

このため本調査研究では、定義語・略称が用いられている条文について、自然言語処理の手法により定義語・略称が定義された箇所を抽出し、それらが使用されている箇所を特定することによって、定義されている条文から同一法令内でそれらが使用されている条文のリンク設定を行う方法、また定義語・略称が使用されている条文からそれらが定義されている条文へのリンク設定がされたユーザー・インターフェース（以下「以下リスト UI」という。）を提案し、それを実現するための調査・実証を行った。

また、定義語・略称を定義する場合、多様な表記方法が存在することで、定義部分を抽出すること、それに加え定義語・略称が使用されている条文がその内容が適用される条件や有効範囲に合致しているかを判定することが単純な文字列のパターンマッチングで処理することは困難であると想定できるため、自然言語処理による解析でこれらのリンク設定の実現可能性について調査・実証し、これらに対する課題類型についての整理分析を行った。

3. 2. 2. 調査方法

調査対象となる法令条文のうち、まずは定義部分と思われる表記パターンの抽出について、これまでの知見から以下に類型化できると仮定した。

- データ構造でパターンマッチング対応できるもの（類型1）
使用されている XML タグ等をヒントに定義部分と判断できるもの。
- 言い回しなど特定の表記的なパターンを持つもの（類型2）
文字列に対する正規表現等で定義部分と判断できるもの。
- 内容の解釈が必要となるもの（類型3）

これらの類型化されたパターンでの自然言語処理を含む機械的な処理による定義部分の抽出が可能かどうか、また抽出した定義語・略語の用いられている箇所を抽出し、定義部分と関連付けする手法についても同様に可能であるかの整理・調査を行うこととした。

また、抽出したデータを使用して「以下リスト UI」のモックアップ¹²（「別添資料 主要プログラムコード」中の「以下リスト UI 参照」）を作成することで、その中で生じる課題も合わせて整理分析を行った。

4. 法令中の併合的接続詞・選択的接続詞の表記に関する調査について整理分析

4. 1. 「条文生成 UI」で新規に成文化する機能の実現可能性について

新規条文の成文化の実現可能性を検討するため、前述の Figma によるプロトタイプを基にして、文章要素を指定した関係に基づく接続詞によって接続し、文を生成する「条文生成 UI」ツール（以降「検証ツール」と呼ぶ。）を作成した。検証ツールの画面例を次に示す。

¹² 開発するソフトウェアの操作感を把握するために作成する簡易な画面表示プログラム。実物と同様の表示や画面遷移が行えるが、一般的には通信等の画面表示以外の機能を持たない

条文生成

代理人の死亡又は代理人が破産手続開始の決定若しくは後見開始の審判を受けたこと。

図 1 0 作成した検証ツール画面

この検証ツールでは、文の生成に関して次の機能を実装している。

1. 要素の関係性（AND/OR/単純結合（画面表記上は「継続」））を指定することで、当該要素間に対応する接続詞により接続する
2. 要素が親子構造になっている場合、要素の位置する段数に応じて接続詞を可変させる（いわゆる「大若し・小若し」、「大並び・小並び」）

このツールを用いて既存条文を実際に生成できるかを評価した結果、次のような課題が得られた。

4. 1. 1. 読点による接続の課題

同一関係の要素が複数個並列に存在する場合、次のルールによって要素を接続する必要がある。

- 名詞が 3 つ以上並ぶ場合、末尾の接続以外は読点のみで連結し、末尾の接続については接続詞で連結する

例)

地上権、永小作権、質権、使用貸借による権利、賃借権若しくはその他の使用及び収益を目的とする権利

- 動詞・形容詞・副詞が2つ以上並ぶ場合、接続詞の前に読点を配置し、また3つ以上並ぶ場合は末尾の接続以外は読点のみで連結する

例)

当該許可に係る掘削の工事を完了し、又は廃止したときは

これらのルールを実現する場合、次のどちらかのアプローチによって適用するルールを判断する必要がある。

1. 品詞解析による要素の品詞の判定
各要素の品詞を判定することで、どのルールが当てはまるかを判断する
2. 要素のパターンマッチングによる処理
要素種別を分類するパターンルールを用いる。例えば、「し」、「とき」、「させ」等で要素が終端される場合、動詞・形容詞・副詞として取り扱う

今回の検証ツールにおいては、この読点による接続のルールは実装を見送った。これは、アプローチ1の場合には品詞を判定するソフトウェアを検証ツールに組み込む必要があり、Webブラウザ上のプログラムでは実現が困難なこと、またアプローチ2の場合には準備するパターンルールの検討に多大な時間を要することが見込まれたためである。

適切な条文の成文化のためには読点による接続への対応は必須と考えられるため、上記のいずれかのアプローチによる実現を検討する必要がある。この際、それぞれの判定精度を評価し、採用するアプローチを検討することが望ましいと思われる。

4. 1. 2. 入力操作の課題

既存条文による検証において、複雑な親子構造を持つ条文については、入力操作が難解になる場合が見受けられた。次の条文を例として挙げる。

第一項の歳入並びに第二項の手数料及び延滞金の還付並びにこれらの徴収金の徴収又は還付に関する書類の送達及び公示送達については、地方税の例による。

(地方自治法（昭和二十二年法律第六十七号）第二百三十一条の三第四項）

この条文を検証ツールに入力する場合、次のような親子構造を画面上に組み上げた上で入力する必要がある。

図 1 1 検証ツール画面の入力例

このような複雑な親子構造を持つ条文の場合、あらかじめ条文の各要素がそれぞれのレベルで接続されるのかを把握しておかなければ正しい入力が困難である。

この点について、検証ツールの操作性の工夫による改善が有効であると思われる。例えば、文の全ての要素を単一のパーツとして登録し、それらをドラッグするなどして適切な位置階層に配置するようなオペレーションが考えられる。ただし、実際の操作方法や画面のデザインの検討、画面の実装作業には相応の時間を要するものと思われる。

4. 2. 「条文生成 UI」で既存条文を扱う実現可能性について

既存条文を「条文生成 UI」で扱う場合、次の判断をプログラムによって行う必要がある。

- 各接続詞が接続対象とする要素及び範囲の特定
- 各接続詞が条文中で属する入れ子階層の特定

これらの実現可能性について検討を行った。

4. 2. 1. 各接続詞が接続対象とする要素及び範囲の特定について

接続対象となる単語の特定については、基本的には接続詞の前後にある要素を抽出することによって行うことができる。ただし、次の場合には要素を特定することが困難となる。

- 要素が単語でない場合
- 読点によって接続されている場合

まず、要素が単語でない場合については、どこまでが要素であるのかをプログラム処理で機械的に判断することは不可能である。例えば次の条文例では、先頭の「第一項の歳入」が一つの要素となるが、プログラム処理では「歳入」のみを要素として判断してしまう。

第一項の歳入並びに第二項の手数料及び延滞金の還付並びにこれらの徴収金の徴収又は還付に関する書類の送達及び公示送達については、地方税の例による。
(地方自治法(昭和二十二年法律第六十七号)第二百三十一条の三第四項)

また、読点によって接続されている場合も、その読点が接続詞の従属としての読点であるのか、または要素自体に読点が含まれているのかをプログラム処理で機械的に判断することは不可能である。

これらの問題については、純粋な条文からは正しい判断を行うことができないため、あらかじめデータ上で範囲を明示するような人的作業が必要となる。

4. 2. 2. 各接続詞が条文中で属する入れ子階層の特定

入れ子階層の特定に関しては、同じ接続詞が複数の階層に出現する場合があるため、機械的な解析は困難である。ここでは、次の例で考える。

管理者は、共用部分並びに第二十一条に規定する場合における当該建物の敷地及び附属施設（次項及び第四十七条第六項において「共用部分等」という。）を保存し、集会の決議を実行し、並びに規約で定めた行為をする権利を有し、義務を負う。
(建物の区分所有等に関する法律（昭和三十七年法律第六十九号）第二十六条第一項)

この場合、1つ目の「並びに」は第2階層に位置し、2つ目の「並びに」は第1階層に位置することとなる。この判断を行うためには、条文の意味を解釈した上で適切な階層を検討する必要があり、機械的な処理で行うことは困難であると思われる。

したがって、各接続詞が接続対象とする要素及び範囲の特定の場合と同様、あらかじめデータに意味を付与するような人的作業が必要となる。

これらの検討により、既存条文を「条文生成 UI」で扱うためには、対象条文に対して人的作業によって構造情報を付与する必要があり、純粋な条文を機械的処理のみで扱うことは、現時点では技術的に困難であると考えられる。

4. 3. データ構造に起因する課題の整理

4. 3. 1. 課題の種類一覧

本調査研究において、併合的接続詞・選択的接続詞を用いた条文は自然言語処理での対応が可能なテキストデータを対象としている。しかしながら分析対象となった法令 XML については、併合的接続詞・選択的接続詞を用いた表現の中には QuoteStruct タグを含んでいるものが存在した。QuoteStruct タグは「改行を含む構造の引用を表す要素」であり、文章中に画像や表組をインライン（行中）に挿入する場合に用いられるが、データ構造上はあくまで文字列と並列のものとして扱われる。つまり文字列以外のものに対する併合的接続詞・選択的接続詞を用いた表現が存在していた。

QuoteStruct タグで表現されるものを併合的接続詞・選択的接続詞で結び付けているものは以下のとおりである。

	パターン	課題
1	Fig タグが使用されている表現	画像タグ同士を接続対象としている
2	TableStruct タグが使用されている表現	表組タグ同士を接続対象としている

4. 3. 2. 課題の類型に関する事例

備考
— この算式において、 $L_{E,di}$ 、 $L_{E,dj}$ 、 $L_{E,dk}$ 、 T_0 及び T の意義は、それぞれ次のとおりとする。

特定空港周辺航空機騒音対策特別措置法施行規則（昭和五十三年運輸省・建設省令第二号）第二条

図12 Fig タグが使用されている表現

(施行期日)			
第一条 この省令は、昭和五十五年五月二十五日（以下「施行日」という。）から施行する。ただし、第一条中目次の改正規定（「／第七編 昇降設備／第八編 コンテナ設備／」を改める部分に限る。）、第七編の編名を改める改正規定、第七編中第三百三条の前に章名を付する改正規定、第八編の編名を削る改正規定、第三百十一條の次に章名を付する改正規定及び第七編に一章を加える改正規定、第十一条中目次の改正規定及び第十一章を第十二章とし、第十章の次に一章を加える改正規定、第十二条中別表第一の改正規定（「			
コンテナ	フラットトラック型のもの	1個につき	11,000円
	その他の型のもの	1個につき	15,000円
」を改める部分に限る。）並びに第十三条中別表の改正規定（「			
コンテナ	フラットトラック型のもの	68,000	1箇につき2,200
	その他の型のもの	98,000	2,800
」を改める部分に限る。）並びに附則第二条第十四項及び附則第十二条第三項の規定は、公布の日から施行する。			
(船舶設備規程の一部改正に伴う経過措置)			

特定空港周辺航空機騒音対策特別措置法施行規則（昭和五十三年運輸省・建設省令第二号）第二条

図13 TableStruct タグが使用されている表現

4. 4. 法令の表記ルールに起因する課題の整理

4. 4. 1. 課題の類型

法令の表記ルールに起因する課題として、次の2点が見受けられた。以降でそれぞれの内容について説明する。

- 読点の判断
- 括弧書きで用いられる接続詞

4. 4. 2. 課題の類型に関する事例

読点の判断

前述のように併合的接続詞・選択的接続詞が用いられる法令条文においては、読点の使用について品詞による判断に基づいたルールが存在する。このルールに従い、併合的接続詞・選択的接続詞と読点を使用し、法令条文を「オブジェクト表現」する場合には、自然言語処理によって文章をトークン（処理・分析するための基本的な構成単位）に分割し、品詞情報を付与しなければならない。ただし、既存条文に品詞情報を付与するにあたっては、法令条文で用いられる固有の用語等への対応を可能にするため、既存条文を法令条文に特化した辞書データを用いた形態素解析を実施し、あらかじめ適切な品詞情報を付与することが必要となる。

また、新規条文を「オブジェクト表現」するためには、条文生成 UI を利用する場合にも文章を入力する際に形態素解析を実施し、適切な品詞情報を付与する必要がある。この自然言語処理には、処理のための専用ソフトウェア実行環境の準備が必要となる。

● 事例サンプル

食品衛生法（昭和二十二年法律第二百三十三号）第十三条第二項

前項の規定により基準又は規格が定められたときは、その基準に合わない方法により食品若しくは添加物を製造し、加工し、使用し、調理し、若しくは保存し、その基準に合わない方法による食品若しくは添加物を販売し、若しくは輸入し、又はその規格に合わない食品若しくは添加物を製造し、輸入し、加工し、使用し、調理し、保存し、若しくは販売してはならない。

括弧書きで用いられる接続詞

法令条文での丸括弧を用いた表現は「定義語や略称を示す」ことや「読み替えや一部改正の前後の字句を示す」ことなど、その文章の大意を補足する形で用いられるが、このような丸括弧を用いた表現中にも併合的接続詞・選択的接続詞は多く見受けられる。可読性を向上させることを目的として、併合的接続詞・選択的接続詞を用いた表現を「オブジェクト表現」することは可能であるものの、それらを多く提示することで可読性を損なうことがあると考えられる。このような法令条文については、ある程度の可読性を担保した形での表現となるような線引きが必要と考える。

● 事例サンプル

銀行法（昭和五十六年法律第二十一号）第二条第十一項

第八項又は前項の場合において、会社又は議決権の保有者が保有する議決権には、金銭又は有価証券の信託に係る信託財産として所有する株式等に係る議決権（委託者又は受益者が行使し、又はその行使について当該会社若しくは当該議決権の保有者に指図を行うことができるものに限る。）その他内閣府令で定める議決権を含まないものとし、信託財産である株式等に係る議決権で、当該会社又は当該議決権の保有者が委託者若しくは受益者として行使し、又はその行使について指図を行うことができるもの（内閣府令で定める議決権を除く。）及び社債、株式等の振替に関する法律（平成十三年法律第七十五号）第四百七条第一項又は第四百八条第一項の規定により発行者に対抗することができない株式に係る議決権を含むものとする。

4. 5. 内容解釈に起因する課題の整理

4. 5. 1. 課題の類型一覧

内容解釈に起因する課題として、次の1点が見受けられた。以降でその内容を説明する。

- 「大若し・小若し」、「大並び・子並び」の判断

4. 5. 2. 課題の類型に関する事例

「大若し・小若し」、「大並び・子並び」の判断

法令条文で用いられる接続詞のうち、「若しくは」と「並びに」は、複数の階層に併存する場合がある。

このような条文について、既存条文の解析を行う場合、どの接続詞がどの階層に位置するかを特定することは機械的に困難である。一方で、新規条文の生成の場合には、正しく「大若し・小若し」、「大並び・子並び」を表現できることが検証ツールによって確認できた。

● 事例サンプル

地方自治法（昭和二十二年法律第六十七号）第二百三十一条の三第四項

第一項の歳入並びに第二項の手数料及び延滞金の還付並びにこれらの徴収金の徴収又は還付に関する書類の送達及び公示送達については、地方税の例による。

4. 6. 例外的な表現について

「かつ」という表現を含むものについて

「及び」、「並びに」と同様の併合的接続詞として「かつ」が使用される場合もある。この場合には「及び」、「並びに」での表記ルール以外のための処理への考慮が必要となる（語ではなく文章を接続する場合には「かつ」の前後に読点を打つというルールも存在する）。

● 事例サンプル

地方自治法（昭和二十二年法律第六十七号）第二条第十一項

地方公共団体に関する法令の規定は、地方自治の本旨に基づき、かつ、国と地方公共団体との適切な役割分担を踏まえたものでなければならない。

「A又はBは、～してはならない」、「A及びBは、～してはならない」の両方の表現があるものについて

「Aも「B」もある行為をしてはならない、というような場合には「A又はBは、～してはならない。」と「A及びBは、～してはならない」の両方の表現があり、現実にはいずれの立法例も存在することから、これらの表現に対する対応も踏まえる必要がある。

● 事例サンプル1

住民基本台帳法（昭和四十二年法律第八十一号）第三十五条

住民基本台帳に関する調査に関する事務に従事している者又は従事していた者は、その事務に関して知り得た秘密を漏らしてはならない。

● 事例サンプル2

地方自治法（昭和二十二年法律第六十七号）第二条第十六項

地方公共団体は、法令に違反してその事務を処理してはならない。なお、市町村及び特別区は、当該都道府県の条例に違反してその事務を処理してはならない

文中で要素同士が対応する表現について

文の前半と後半にそれぞれ接続詞による並列表現が存在し、前半の各要素と後半の各要素とが対応付けられている表現がある。これを入力補助的な観点から考えると、紐づく要素のペアを視覚化できるような機能などを検討する余地があるといえる。

- 事例サンプル

労働委員会規則（昭和二十四年中央労働委員会規則第一号）第二条第一項

「労組法」、「労調法」、「行労法」、「地方公労法」、「労組法施行令」、「労調法施行令」及び「行労法施行令」とは、それぞれ労働組合法、労働関係調整法、行政執行法人の労働関係に関する法律、地方公営企業等の労働関係に関する法律、労働組合法施行令（昭和二十四年政令第二百三十一号）、労働関係調整法施行令（昭和二十一年勅令第四百七十八号）及び行政執行法人の労働関係に関する法律施行令（昭和三十一年政令第二百四十九号）をいう。

4. 7. 調査結果に基づく考察

新規条文の成文化については、読点への対応を除いた形での検証ツールの実装を行うことができた。今後の課題として次の事項が挙げられる。

- 読点での接続への対応

読点での接続が必要であるかを判断するためには、対象要素の品詞を判定する必要がある。この判定処理の計算量は膨大となるため、実装に当たっては処理の実行環境について考慮が必要になると考えられる。また、判定の精度を高めるため、法令条文に特化したコーパス¹³の整備があることが望ましい。

- 入力時の構造理解の必要性

本調査研究で作成した検証ツールにおいては、あらかじめ入力者が各要素の入れ子構造を理解していないと入力操作が難しいものとなってしまった。しかしながら、入れ子構造を理解している場合には、今回の検証ツールのような入力補助機能の有効性がそれほど大きくなるとも考えられる。このため、入力補助機能の利便性は、入力者の入れ子構造の理解を補助するよりも、入力者が入れ子構造を可視化する作業を効率化する点に重点を置いて意義付けることが考えられる。入れ子構造をより簡易な形での入力操作を実現するには、ユーザー・インターフェースの検討が必要になると考えられる。

既存条文の解析については、本調査研究においては各要素の示す範囲や入れ子構造における位置特定の難しさから、現時点では機械的処理は困難であるとの結論に至った。しかしながら、パターンマッチングによる各要素の分割処理を実施することで、文章構造の視覚化を実現

¹³ 機械的な処理のため、自然言語の文章に言語的な情報を付与して構造化し、大規模に蓄積したものである。

できる可能性がある。例えば、読点や接続詞、助詞によって文章中の文字列を分割し、画面上にそれぞれを部品として提示するようなアプローチである。

この「読点や接続詞、助詞によって文章中の文字列を分割し、画面上にそれぞれを部品として提示するようなアプローチ」の提案に至った要因のひとつとして、分析時に作成した文章構造把握のためのツールの存在がある。これによって、その文章でどのような単語が用いられているのかを俯瞰的に把握することができると思われる。また、誤った分割がなされた場合には、実際の条文にも誤読のリスクを含み得るとも考えられるため、そのようなリスクの調査の観点からも有用といえる。

将来の取組に当たっては、このようなアプローチの検討も意義があると考えている。

5. 法令中の定義語・略称表記に関する調査についての整理分析

5. 1. 「以下リスト UI」の実現可能性について

今回、定義語・略称が用いられている条文について、それらの自然言語処理による抽出・特定及び定義箇所・使用箇所へのリンクを生成する検証を行った。この検証において作成したリンク設定済条文表示画面を「以下リスト UI」と呼ぶ。

検証に際しては、次の流れ（ステップ）で以下リスト UI の作成を行った。

1. 自然言語処理による定義語・略称の識別が可能な表記の種類（パターン）を特定
2. それらの様式を用いている法令を、各々1~2つサンプリング
3. サンプリングした法令について、自然言語処理によって以下リスト UI の基となるデータを生成し、以下リスト UI を実装

以降では、上記各ステップにおける対応・検証内容を説明する。

ステップ1：自然言語処理による定義語・略称の識別が可能な表記の種類（パターン）の特定

定義語・略称の表記パターンの分類のため、本書冒頭の「調査対象」に挙げた 8,956 法令について、見出しに「定義」、「定義等」、「用語の定義」、「～の定義」、「用語の定義及び略称」、「この規則の趣旨・定義」、「定義及び～」、「～の定義に関する事項」を含む条文を定義条文とし（対象法令：2,209 法令）、それらに含まれる表記パターンの調査を行った。今回の調査研究ではパターンごとの件数の全量調査をしてはいないが、複数の事例が見られたものを類型化している。機械的な処理の実現可能性の観点より、次の 5 つに表記パターンを類型化した。

類型	「3. 2. 2. 調査方法」で仮定した類型	内容	説明
パターン A	類型 1（構造ベース）	自然言語処理による抽出と XML データ構造を利用して処理できるもの	列挙パターン
パターン B	類型 1（構造ベース） + 類型 2（表記パターン）	多少複雑な自然言語処理が必要となるもの	<ul style="list-style-type: none"> かぎ括弧での文章パターン（ただし書きなどもある） かぎ括弧なしの文章パターン（略称の定義も含まれる場合あり）
パターン C	類型 2（表記パターン） + 類型 3（意味解釈）	複雑な自然言語処理や法制執務的な観点からのメタ情報が必要となるもの	<ul style="list-style-type: none"> 定義条文等で定義されるが、その後、別の場所で再定義されるようなもの 先に定義語が出現し、その後その定義がされるようなもの 条件が先立つパターン 他の法令での定義を使用するパターン
パターン D	類型 3（意味解釈）	表記方法や内容解釈を理由として対応するハードルが高いもの	<ul style="list-style-type: none"> 除外条件があるもの（「～条は除く」のように「出現位置のような文脈的な除外」と捉えられるものと「～する者に限る」、「～の場合は除く」のように「法解釈上での除外」と捉えられるもの）

			<ul style="list-style-type: none"> 括弧書きの入れ子構造や並列配置などの表現（定義中の定義、定義と有効範囲を同じ階層で、定義と有効範囲を並列の括弧書き等）
パターン E		これまでの分析に加えてさらなる分析が必要であるもの	見出しなどで定義が含まれていることが示されていないもの

ステップ 2：パターンごとに法令をサンプリング

上記の分類において、該当する法令について分類を行った。

類型	法令名	特徴
パターン A	酒税法（昭和二十八年法律第六号）第三条	列挙パターン
	金融商品取引法（昭和三十二年法律第二十五号）第二条	
パターン B	民法（明治二十九年法律第八十九号）第八十五条	文章内でかぎ括弧が使用されて定義されるパターン
	職業安定法施行規則（昭和二十二年労働省令第十二号）第一条	文章内でかぎ括弧が使用されずに定義されるパターン
	労働委員会規則（昭和二十四年中央労働委員会規則第一号）第二条	読点で列挙された括弧書きを対応させる必要がある
パターン C	商法（明治三十二年法律第四十八号）第四条と第十一条	定義条文等で定義されるがその後再定義される
	少年警察活動規則（平成十四年国家公安委員会規則第二十号）第六条・第七条第一項	先に定義語が出現しその後に定義部分が出現する
	協同組合による金融事業に関する法律（昭和二十四年法律第百八十三号）第四条第一項	条件が先立つ
	特定計量器検定検査規則（平成五年通商産業省令第七十号）第一条	他の法令での定義を使用する
パターン D		定義に関する除外条件かどうか判断できないもの

<p>地方公共団体の手数料の標準に関する政令に規定する総務省令で定める金額等を定める省令（平成十二年自治省令第五号）第二条</p> <p>地方税法第五十三条</p>	<p>括弧書きが入れ子構造になっていたり並列配置されたりしていて表記的な側面からは判断できないものの</p>
--	--

これらのうち、課題を持つものについて詳細は後述する。

ステップ3：「以下リスト UI」生成用データの作成と実装

パターン C については、高度な自然言語処理（法令条文に特化した品詞情報の付与や係り受け分析等）、パターン D については現在の表記方法の見直しや内容解釈が必要であることから、現時点で「自然言語処理を含む機械的な処理での生成」が不可であると判断し、機械的な処理のみで生成が可能と思われる以下の2パターンについて「以下リスト UI」の基となるデータ作成と実装を行った。

パターン A

- 酒税法
- 金融商品取引法

パターン B

- 民法（第四章のみ）
- 職業安定法施行規則
- 労働委員会規則

「以下リスト UI」生成用データは、定義語・略語が定義・参照されている箇所を一意で特定するために、e-Gov 法令検索で使用されている HTML データから、定義部分と思われる箇所と使用箇所、及びリンク情報の付与のために HTML タグの ID を機械処理で抽出した。この後、抽出内容を人手によりチェックし Excel データとして作成した。

	A	B	C	D	E	F
1	初回出荷場所	規定された品	定義	指示品項	備考メモ	
7	第三系第1項第4号	醸造酒類+Mp-Ch_1-At_3-Pr_1-Ih_4	次に掲げる酒類（その他の発泡性酒類を除く。） イ 清酒 ロ 果実酒 ハ その他の醸造酒	第二十三系第1項第2号+Mp-Ch_3-At_23-Pr_1-Ih_2 第二十三系第2項にも規定された品があります		
8	第三系第1項第5号	高糖酒類+Mp-Ch_1-At_3-Pr_1-Ih_5	次に掲げる酒類（その他の発泡性酒類を除く。） イ 濃縮式高糖酒類 ロ 醸造酒類 ハ ウイスキー ニ ブランデー ホ 原料用アルコール ヘ スピリッツ	第二十三系第1項第3号+Mp-Ch_3-At_23-Pr_1-Ih_3 第二十三系第2項+Mp-Ch_3-At_23-Pr_3	第二系第2項にも規定された品があります	
9	第三系第1項第6号	醸成酒類+Mp-Ch_1-At_3-Pr_1-Ih_6	次に掲げる酒類（その他の発泡性酒類を除く。） イ 合成果酒 ロ みりん ハ 甘味果実酒 ニ リキュール ホ 紹興酒 ヘ 雑酒	第二十三系第1項第4号+Mp-Ch_3-At_23-Pr_1-Ih_4 第二十三系第4項+Mp-Ch_3-At_23-Pr_4	第二系第2項にも規定された品があります	
10	第三系第1項第7号	清酒+Mp-Ch_1-At_3-Pr_1-Ih_7	次に掲げる酒類でアルコール分が二十度未満のものを含む。 イ 米、米こうじ及び水を原料として発酵させて、こしたものを、 ロ 米、米こうじ、水及び清酒かすその他成分で定められた原料として発酵させて、こしたものを、 ハ 酒類に清酒かすを加えて、こしたものを、 ニ 酒類に清酒かすを加えて、こしたものを、 ホ 酒類に清酒かすを加えて、こしたものを、 ヘ 酒類に清酒かすを加えて、こしたものを、 （その原料中当該成分で定められた物品の重量の合計が米（こし米を含む。）の重量の百分の五を超過するものを含む。） ハ 酒類に清酒かすを加えて、こしたものを、 ニ 酒類に清酒かすを加えて、こしたものを、 ホ 酒類に清酒かすを加えて、こしたものを、 ヘ 酒類に清酒かすを加えて、こしたものを、 （その原料中当該成分で定められた物品の重量の合計が米（こし米を含む。）の重量の百分の五を超過するものを含む。）	第三系第1項第7号+Mp-Ch_1-At_3-Pr_1-Ih_7 第三系第1項第8号+Mp-Ch_1-At_3-Pr_1-Ih_8 第七系第2項第1号+Mp-Ch_2-At_7-Pr_2-Ih_1 第七系第3項第1号+Mp-Ch_2-At_7-Pr_3-Ih_1 第七系第3項第2号+Mp-Ch_2-At_7-Pr_3-Ih_2 第七系第3項第3号+Mp-Ch_2-At_7-Pr_3-Ih_3 第七系第3項第4号+Mp-Ch_2-At_7-Pr_3-Ih_4 第七系第3項第5号+Mp-Ch_2-At_7-Pr_3-Ih_5 第七系第3項第6号+Mp-Ch_2-At_7-Pr_3-Ih_6 第七系第3項第7号+Mp-Ch_2-At_7-Pr_3-Ih_7 第七系第3項第8号+Mp-Ch_2-At_7-Pr_3-Ih_8 第七系第3項第9号+Mp-Ch_2-At_7-Pr_3-Ih_9 第七系第3項第10号+Mp-Ch_2-At_7-Pr_3-Ih_10 第七系第3項第11号+Mp-Ch_2-At_7-Pr_3-Ih_11 第七系第3項第12号+Mp-Ch_2-At_7-Pr_3-Ih_12 第七系第3項第13号+Mp-Ch_2-At_7-Pr_3-Ih_13 第七系第3項第14号+Mp-Ch_2-At_7-Pr_3-Ih_14 第七系第3項第15号+Mp-Ch_2-At_7-Pr_3-Ih_15 第七系第3項第16号+Mp-Ch_2-At_7-Pr_3-Ih_16 第七系第3項第17号+Mp-Ch_2-At_7-Pr_3-Ih_17 第七系第3項第18号+Mp-Ch_2-At_7-Pr_3-Ih_18 第七系第3項第19号+Mp-Ch_2-At_7-Pr_3-Ih_19 第七系第3項第20号+Mp-Ch_2-At_7-Pr_3-Ih_20 第七系第3項第21号+Mp-Ch_2-At_7-Pr_3-Ih_21 第七系第3項第22号+Mp-Ch_2-At_7-Pr_3-Ih_22 第七系第3項第23号+Mp-Ch_2-At_7-Pr_3-Ih_23 第七系第3項第24号+Mp-Ch_2-At_7-Pr_3-Ih_24 第七系第3項第25号+Mp-Ch_2-At_7-Pr_3-Ih_25 第七系第3項第26号+Mp-Ch_2-At_7-Pr_3-Ih_26 第七系第3項第27号+Mp-Ch_2-At_7-Pr_3-Ih_27 第七系第3項第28号+Mp-Ch_2-At_7-Pr_3-Ih_28 第七系第3項第29号+Mp-Ch_2-At_7-Pr_3-Ih_29 第七系第3項第30号+Mp-Ch_2-At_7-Pr_3-Ih_30 第七系第3項第31号+Mp-Ch_2-At_7-Pr_3-Ih_31 第七系第3項第32号+Mp-Ch_2-At_7-Pr_3-Ih_32 第七系第3項第33号+Mp-Ch_2-At_7-Pr_3-Ih_33 第七系第3項第34号+Mp-Ch_2-At_7-Pr_3-Ih_34 第七系第3項第35号+Mp-Ch_2-At_7-Pr_3-Ih_35 第七系第3項第36号+Mp-Ch_2-At_7-Pr_3-Ih_36 第七系第3項第37号+Mp-Ch_2-At_7-Pr_3-Ih_37 第七系第3項第38号+Mp-Ch_2-At_7-Pr_3-Ih_38 第七系第3項第39号+Mp-Ch_2-At_7-Pr_3-Ih_39 第七系第3項第40号+Mp-Ch_2-At_7-Pr_3-Ih_40 第七系第3項第41号+Mp-Ch_2-At_7-Pr_3-Ih_41 第七系第3項第42号+Mp-Ch_2-At_7-Pr_3-Ih_42 第七系第3項第43号+Mp-Ch_2-At_7-Pr_3-Ih_43 第七系第3項第44号+Mp-Ch_2-At_7-Pr_3-Ih_44 第七系第3項第45号+Mp-Ch_2-At_7-Pr_3-Ih_45 第七系第3項第46号+Mp-Ch_2-At_7-Pr_3-Ih_46 第七系第3項第47号+Mp-Ch_2-At_7-Pr_3-Ih_47 第七系第3項第48号+Mp-Ch_2-At_7-Pr_3-Ih_48 第七系第3項第49号+Mp-Ch_2-At_7-Pr_3-Ih_49 第七系第3項第50号+Mp-Ch_2-At_7-Pr_3-Ih_50 第七系第3項第51号+Mp-Ch_2-At_7-Pr_3-Ih_51 第七系第3項第52号+Mp-Ch_2-At_7-Pr_3-Ih_52 第七系第3項第53号+Mp-Ch_2-At_7-Pr_3-Ih_53 第七系第3項第54号+Mp-Ch_2-At_7-Pr_3-Ih_54 第七系第3項第55号+Mp-Ch_2-At_7-Pr_3-Ih_55 第七系第3項第56号+Mp-Ch_2-At_7-Pr_3-Ih_56 第七系第3項第57号+Mp-Ch_2-At_7-Pr_3-Ih_57 第七系第3項第58号+Mp-Ch_2-At_7-Pr_3-Ih_58 第七系第3項第59号+Mp-Ch_2-At_7-Pr_3-Ih_59 第七系第3項第60号+Mp-Ch_2-At_7-Pr_3-Ih_60 第七系第3項第61号+Mp-Ch_2-At_7-Pr_3-Ih_61 第七系第3項第62号+Mp-Ch_2-At_7-Pr_3-Ih_62 第七系第3項第63号+Mp-Ch_2-At_7-Pr_3-Ih_63 第七系第3項第64号+Mp-Ch_2-At_7-Pr_3-Ih_64 第七系第3項第65号+Mp-Ch_2-At_7-Pr_3-Ih_65 第七系第3項第66号+Mp-Ch_2-At_7-Pr_3-Ih_66 第七系第3項第67号+Mp-Ch_2-At_7-Pr_3-Ih_67 第七系第3項第68号+Mp-Ch_2-At_7-Pr_3-Ih_68 第七系第3項第69号+Mp-Ch_2-At_7-Pr_3-Ih_69 第七系第3項第70号+Mp-Ch_2-At_7-Pr_3-Ih_70 第七系第3項第71号+Mp-Ch_2-At_7-Pr_3-Ih_71 第七系第3項第72号+Mp-Ch_2-At_7-Pr_3-Ih_72 第七系第3項第73号+Mp-Ch_2-At_7-Pr_3-Ih_73 第七系第3項第74号+Mp-Ch_2-At_7-Pr_3-Ih_74 第七系第3項第75号+Mp-Ch_2-At_7-Pr_3-Ih_75 第七系第3項第76号+Mp-Ch_2-At_7-Pr_3-Ih_76 第七系第3項第77号+Mp-Ch_2-At_7-Pr_3-Ih_77 第七系第3項第78号+Mp-Ch_2-At_7-Pr_3-Ih_78 第七系第3項第79号+Mp-Ch_2-At_7-Pr_3-Ih_79 第七系第3項第80号+Mp-Ch_2-At_7-Pr_3-Ih_80 第七系第3項第81号+Mp-Ch_2-At_7-Pr_3-Ih_81 第七系第3項第82号+Mp-Ch_2-At_7-Pr_3-Ih_82 第七系第3項第83号+Mp-Ch_2-At_7-Pr_3-Ih_83 第七系第3項第84号+Mp-Ch_2-At_7-Pr_3-Ih_84 第七系第3項第85号+Mp-Ch_2-At_7-Pr_3-Ih_85 第七系第3項第86号+Mp-Ch_2-At_7-Pr_3-Ih_86 第七系第3項第87号+Mp-Ch_2-At_7-Pr_3-Ih_87 第七系第3項第88号+Mp-Ch_2-At_7-Pr_3-Ih_88 第七系第3項第89号+Mp-Ch_2-At_7-Pr_3-Ih_89 第七系第3項第90号+Mp-Ch_2-At_7-Pr_3-Ih_90 第七系第3項第91号+Mp-Ch_2-At_7-Pr_3-Ih_91 第七系第3項第92号+Mp-Ch_2-At_7-Pr_3-Ih_92 第七系第3項第93号+Mp-Ch_2-At_7-Pr_3-Ih_93 第七系第3項第94号+Mp-Ch_2-At_7-Pr_3-Ih_94 第七系第3項第95号+Mp-Ch_2-At_7-Pr_3-Ih_95 第七系第3項第96号+Mp-Ch_2-At_7-Pr_3-Ih_96 第七系第3項第97号+Mp-Ch_2-At_7-Pr_3-Ih_97 第七系第3項第98号+Mp-Ch_2-At_7-Pr_3-Ih_98 第七系第3項第99号+Mp-Ch_2-At_7-Pr_3-Ih_99 第七系第3項第100号+Mp-Ch_2-At_7-Pr_3-Ih_100	第三系第1項第4号イにも規定された品があります	
11	第三系第1項第8号	合成清酒+Mp-Ch_1-At_3-Pr_1-Ih_8	アルコール（次項の規定（アルコール分に関する規定を除く。）に該当する酒類（水以外の物品を加えたものを除く。）でアルコール分が三十九度以上四十五度以下のものを含む。第十四号（以下同じ。）の原料（濃縮式高糖酒類又は醸造酒類）を除く。水以外の物品を加えたものを除く。第十一号において同じ。）又は清酒とたとう糖その他の成分で定められた原料として製造した酒類（当該酒類の原料として米又は米を原料の全部若しくは一部として製造した物品を使用したものについては、米（米を原料の全部又は一部として製造した物品の原料となつた米を含む。）の重量の合計が、アルコール分二十度を超過した場合は当該酒類の重量の百分の五を超過するものを含む。）で、その意味、色若しその他の性状が清酒に類似するもの（アルコール分が十九度未満でアルコール分が五度以上であることその他の成分で定められる条件を	第七系第2項第2号+Mp-Ch_2-At_7-Pr_2-Ih_2 第七系第2項第3号+Mp-Ch_2-At_7-Pr_3-Ih_3 第七系第2項第4号+Mp-Ch_2-At_7-Pr_4-Ih_4 第七系第2項第5号+Mp-Ch_2-At_7-Pr_5-Ih_5 第七系第2項第6号+Mp-Ch_2-At_7-Pr_6-Ih_6 第七系第2項第7号+Mp-Ch_2-At_7-Pr_7-Ih_7 第七系第2項第8号+Mp-Ch_2-At_7-Pr_8-Ih_8 第七系第2項第9号+Mp-Ch_2-At_7-Pr_9-Ih_9 第七系第2項第10号+Mp-Ch_2-At_7-Pr_10-Ih_10 第七系第2項第11号+Mp-Ch_2-At_7-Pr_11-Ih_11 第七系第2項第12号+Mp-Ch_2-At_7-Pr_12-Ih_12 第七系第2項第13号+Mp-Ch_2-At_7-Pr_13-Ih_13 第七系第2項第14号+Mp-Ch_2-At_7-Pr_14-Ih_14 第七系第2項第15号+Mp-Ch_2-At_7-Pr_15-Ih_15 第七系第2項第16号+Mp-Ch_2-At_7-Pr_16-Ih_16 第七系第2項第17号+Mp-Ch_2-At_7-Pr_17-Ih_17 第七系第2項第18号+Mp-Ch_2-At_7-Pr_18-Ih_18 第七系第2項第19号+Mp-Ch_2-At_7-Pr_19-Ih_19 第七系第2項第20号+Mp-Ch_2-At_7-Pr_20-Ih_20 第七系第2項第21号+Mp-Ch_2-At_7-Pr_21-Ih_21 第七系第2項第22号+Mp-Ch_2-At_7-Pr_22-Ih_22 第七系第2項第23号+Mp-Ch_2-At_7-Pr_23-Ih_23 第七系第2項第24号+Mp-Ch_2-At_7-Pr_24-Ih_24 第七系第2項第25号+Mp-Ch_2-At_7-Pr_25-Ih_25 第七系第2項第26号+Mp-Ch_2-At_7-Pr_26-Ih_26 第七系第2項第27号+Mp-Ch_2-At_7-Pr_27-Ih_27 第七系第2項第28号+Mp-Ch_2-At_7-Pr_28-Ih_28 第七系第2項第29号+Mp-Ch_2-At_7-Pr_29-Ih_29 第七系第2項第30号+Mp-Ch_2-At_7-Pr_30-Ih_30 第七系第2項第31号+Mp-Ch_2-At_7-Pr_31-Ih_31 第七系第2項第32号+Mp-Ch_2-At_7-Pr_32-Ih_32 第七系第2項第33号+Mp-Ch_2-At_7-Pr_33-Ih_33 第七系第2項第34号+Mp-Ch_2-At_7-Pr_34-Ih_34 第七系第2項第35号+Mp-Ch_2-At_7-Pr_35-Ih_35 第七系第2項第36号+Mp-Ch_2-At_7-Pr_36-Ih_36 第七系第2項第37号+Mp-Ch_2-At_7-Pr_37-Ih_37 第七系第2項第38号+Mp-Ch_2-At_7-Pr_38-Ih_38 第七系第2項第39号+Mp-Ch_2-At_7-Pr_39-Ih_39 第七系第2項第40号+Mp-Ch_2-At_7-Pr_40-Ih_40 第七系第2項第41号+Mp-Ch_2-At_7-Pr_41-Ih_41 第七系第2項第42号+Mp-Ch_2-At_7-Pr_42-Ih_42 第七系第2項第43号+Mp-Ch_2-At_7-Pr_43-Ih_43 第七系第2項第44号+Mp-Ch_2-At_7-Pr_44-Ih_44 第七系第2項第45号+Mp-Ch_2-At_7-Pr_45-Ih_45 第七系第2項第46号+Mp-Ch_2-At_7-Pr_46-Ih_46 第七系第2項第47号+Mp-Ch_2-At_7-Pr_47-Ih_47 第七系第2項第48号+Mp-Ch_2-At_7-Pr_48-Ih_48 第七系第2項第49号+Mp-Ch_2-At_7-Pr_49-Ih_49 第七系第2項第50号+Mp-Ch_2-At_7-Pr_50-Ih_50 第七系第2項第51号+Mp-Ch_2-At_7-Pr_51-Ih_51 第七系第2項第52号+Mp-Ch_2-At_7-Pr_52-Ih_52 第七系第2項第53号+Mp-Ch_2-At_7-Pr_53-Ih_53 第七系第2項第54号+Mp-Ch_2-At_7-Pr_54-Ih_54 第七系第2項第55号+Mp-Ch_2-At_7-Pr_55-Ih_55 第七系第2項第56号+Mp-Ch_2-At_7-Pr_56-Ih_56 第七系第2項第57号+Mp-Ch_2-At_7-Pr_57-Ih_57 第七系第2項第58号+Mp-Ch_2-At_7-Pr_58-Ih_58 第七系第2項第59号+Mp-Ch_2-At_7-Pr_59-Ih_59 第七系第2項第60号+Mp-Ch_2-At_7-Pr_60-Ih_60 第七系第2項第61号+Mp-Ch_2-At_7-Pr_61-Ih_61 第七系第2項第62号+Mp-Ch_2-At_7-Pr_62-Ih_62 第七系第2項第63号+Mp-Ch_2-At_7-Pr_63-Ih_63 第七系第2項第64号+Mp-Ch_2-At_7-Pr_64-Ih_64 第七系第2項第65号+Mp-Ch_2-At_7-Pr_65-Ih_65 第七系第2項第66号+Mp-Ch_2-At_7-Pr_66-Ih_66 第七系第2項第67号+Mp-Ch_2-At_7-Pr_67-Ih_67 第七系第2項第68号+Mp-Ch_2-At_7-Pr_68-Ih_68 第七系第2項第69号+Mp-Ch_2-At_7-Pr_69-Ih_69 第七系第2項第70号+Mp-Ch_2-At_7-Pr_70-Ih_70 第七系第2項第71号+Mp-Ch_2-At_7-Pr_71-Ih_71 第七系第2項第72号+Mp-Ch_2-At_7-Pr_72-Ih_72 第七系第2項第73号+Mp-Ch_2-At_7-Pr_73-Ih_73 第七系第2項第74号+Mp-Ch_2-At_7-Pr_74-Ih_74 第七系第2項第75号+Mp-Ch_2-At_7-Pr_75-Ih_75 第七系第2項第76号+Mp-Ch_2-At_7-Pr_76-Ih_76 第七系第2項第77号+Mp-Ch_2-At_7-Pr_77-Ih_77 第七系第2項第78号+Mp-Ch_2-At_7-Pr_78-Ih_78 第七系第2項第79号+Mp-Ch_2-At_7-Pr_79-Ih_79 第七系第2項第80号+Mp-Ch_2-At_7-Pr_80-Ih_80 第七系第2項第81号+Mp-Ch_2-At_7-Pr_81-Ih_81 第七系第2項第82号+Mp-Ch_2-At_7-Pr_82-Ih_82 第七系第2項第83号+Mp-Ch_2-At_7-Pr_83-Ih_83 第七系第2項第84号+Mp-Ch_2-At_7-Pr_84-Ih_84 第七系第2項第85号+Mp-Ch_2-At_7-Pr_85-Ih_85 第七系第2項第86号+Mp-Ch_2-At_7-Pr_86-Ih_86 第七系第2項第87号+Mp-Ch_2-At_7-Pr_87-Ih_87 第七系第2項第88号+Mp-Ch_2-At_7-Pr_88-Ih_88 第七系第2項第89号+Mp-Ch_2-At_7-Pr_89-Ih_89 第七系第2項第90号+Mp-Ch_2-At_7-Pr_90-Ih_90 第七系第2項第91号+Mp-Ch_2-At_7-Pr_91-Ih_91 第七系第2項第92号+Mp-Ch_2-At_7-Pr_92-Ih_92 第七系第2項第93号+Mp-Ch_2-At_7-Pr_93-Ih_93 第七系第2項第94号+Mp-Ch_2-At_7-Pr_94-Ih_94 第七系第2項第95号+Mp-Ch_2-At_7-Pr_95-Ih_95 第七系第2項第96号+Mp-Ch_2-At_7-Pr_96-Ih_96 第七系第2項第97号+Mp-Ch_2-At_7-Pr_97-Ih_97 第七系第2項第98号+Mp-Ch_2-At_7-Pr_98-Ih_98 第七系第2項第99号+Mp-Ch_2-At_7-Pr_99-Ih_99 第七系第2項第100号+Mp-Ch_2-At_7-Pr_100-Ih_100		

図 1 4 「以下リスト UI」生成用 Excel データ

上記で作成した「以下リスト UI」生成用データは、別途プログラム処理によりブラウザ内で利用できる構造化データ (JSON) へ変換し、リンク情報を付与した条文を生成する処理を介して「以下リスト UI」を構築する環境を作成した。



図 1 5 「酒税法」以下リスト UI

5. 2. 「以下リスト UI」の機能実現のために自然言語処理が活用できる範囲について

5. 2. 1. 定義条文におけるパターン分類

上記のパターン A 及びパターン B に分類されるものについては、それぞれ「列挙パターン」、「かぎ括弧表記パターン」として、「以下リスト UI」を構築する上で、最終的な人手による確認作業は必要となったものの、定義部分・使用箇所の抽出及びリンク情報の付与について自然言語処理を活用できることが確認できた。

● 「列挙パターン」事例サンプル

酒税法（昭和二十八年法律第六号）第三条

（その他の用語の定義）

第三条 この法律において、次の各号に掲げる用語の意義は、当該各号に定めるところによる。

- 一 アルコール分 温度十五度の時において原容量百分中に含有するエチルアルコールの容量をいう。
- 二 エキス分 温度十五度の時において原容量百立方センチメートル中に含有する不揮発性成分のグラム数をいう。
- 三 発泡性酒類 次に掲げる酒類をいう。
 - イ ビール
 - ロ 発泡酒
 - ハ イ及びロに掲げる酒類以外の酒類で発泡性を有するもの（アルコール分が十一度未満のものに限る。以下「その他の発泡性酒類」という。）

金融商品取引法（昭和二十三年法律第二十五号）第二条

（定義）

第二条 この法律において「有価証券」とは、次に掲げるものをいう。

- 一 国債証券
- 二 地方債証券
- 三 特別の法律により法人の発行する債券（次号及び第十一号に掲げるものを除く。）
- 四 資産の流動化に関する法律（平成十年法律第百五号）に規定する特定社債券
- 五 社債券（相互会社の社債券を含む。以下同じ。）
- 六 特別の法律により設立された法人の発行する出資証券（次号、第八号及び第十一号に掲げるものを除く。）
- 七 協同組織金融機関の優先出資に関する法律（平成五年法律第四十四号。以下「優先出資法」という。）に規定する優先出資証券
- 八 資産の流動化に関する法律に規定する優先出資証券又は新優先出資引受権を表示する証券
- 九～二十一 （略）

「列挙パターン」とは、「酒税法」サンプルで見られる「定義語」と「定義文」が空白文字で連結されるような「文字列のパターンマッチングで定義条文と判別できる」ものや、「金融商品取引法」サンプルのように「定義語」に続いて定義内容が列挙され、その列挙されているものが法令 XML のデータ構造で把握できるものを指す。

これらの「列挙パターン」に分類したものは、定義条文であることを判断するパターンについてさらなる網羅的な調査が必要ではあるものの、定義語と定義文の表記パターンを文字列のパターンマッチングと自然言語処理により抽出できる可能性が高いものも多く見受けられた。

5. 2. 2. 略語表記におけるパターン分類

酒税法（昭和二十八年法律第六号）第六条

（納税義務者）

第六条 酒類の製造者は、その製造場から移出した酒類につき、酒税を納める義務がある。

2 酒類を保税地域から引き取る者（以下「酒類引取者」という。）は、その引き取る酒類につき、酒税を納める義務がある。

上記サンプルの「酒類を保税地域から引き取る者（以下「酒類引取者」という。）」というように丸括弧とかぎ括弧を使用したものが略語表記において頻出するパターンであった。このようなパターンについては、法令条文に特化した辞書の充実が大前提であるものの、それを利用した自然言語処理による品詞情報の付与と係り受け解析によって「略語」にあたる部分と「その略語が指す実体」を抽出する対応も可能であると考えられる。

酒税法第九条

（酒類の販売業免許）

第九条 酒類の販売業又は販売の代理業若しくは媒介業（以下「販売業」と総称する。）をしようとする者は、政令で定める手続により、販売場（継続して販売業をする場所をいう。以下同じ。）ごとにその販売場の所在地（販売場を設けない場合には、住所地）の所轄税務署長の免許（以下「販売業免許」という。）を受けなければならない。ただし、酒類製造者がその製造免許を受けた製造場においてする酒類（当該製造場について第七条第一項の規定により製造免許を受けた酒類と同一の品目の酒類及び第四十四条第一項の承認を受けた酒類に限る。）の販売業及び酒場、料理店その他酒類をもつばら自己の営業場において飲用に供する業については、この限りでない。

上記のパターンの「販売場（継続して販売業をする場所をいう。以下同じ。）」については、丸括弧のみの使用でかぎ括弧は使用されていないものの、このパターンに限らず「丸括弧中に略語表現が存在する」ルールに基づいた網羅的な調査を実施することでパターンを増やし、機械的な処理で対応することが想定できる。

しかしながら、丸括弧とかぎ括弧を使用した略語表記のなかではその略語が有効な「条件」や「範囲」を併記するものもあり、こういったパターンの場合には上記で述べたような「法令条文に特化した辞書を利用した自然言語処理」での抽出が困難だけでなく、その「条件」や

「範囲」に合致する略語の使用箇所を特定し、リンク情報を付与することにおいても機械的な処理で実施することは難しいと考えられる。

酒税法第六条の三

(移出又は引取り等とみなす場合)

第六条の三 次の各号のいずれかに該当するときは、その該当することとなつた時に当該酒類又は酒母若しくはもろみ（以下この条及び第十条第三号において「酒類等」という。）をその製造場から移出したものとみなす。ただし、第四号の場合において、第二十八条第一項の規定の適用を受けて酒類の製造場から移出する当該酒類については、この限りでない。

上記の例では「酒類又は酒母若しくはもろみ（以下この条及び第十条第三号において「酒類等」という。）」が略語表記の定義部分であるが、「この条」と「第十条第三号」で使用されている箇所のみを略語の有効範囲としなければならない。加えて「この条」という表現においても「第六条の三」ということを特定する必要があり、品詞情報の付与や係り受け解析のような自然言語処理以外のもので情報を取得しなければならないことを確認できた。

5. 3. 法令の表記ルールに起因する課題の整理

5. 3. 1. 課題の類型一覧

本調査研究において、「以下リスト UI」を実現するための基データを生成するにあたり、法令の表記ルールに起因する課題として、以下のものが見受けられた。以降でそれぞれの内容について説明する。

- 読点で連結された「かぎ括弧書き」で定義するもの
- 定義条文等で定義されるがその後再定義されるもの
- 先に定義語が出現しその後定義文が現れるもの
- 定義語・定義文よりも先立って条件が来るもの
- 他の法令での定義を使用するもの

5. 3. 2. 課題の類型に関する事例

読点で連結された「かぎ括弧書き」で定義するもの

労働委員会規則（昭和二十四年中央労働委員会規則第一号）第二条

（用語の定義及び略称）

第二条 この規則中次に掲げる用語は、別段の定めがある場合を除き、それぞれ次の意味に用いる。

一 「労組法」、「労調法」、「行労法」、「地方公労法」、「労組法施行令」、「労調法施行令」及び「行労法施行令」とは、それぞれ労働組合法、労働関係調整法、行政執行法人の労働関係に関する法律、地方公営企業等の労働関係に関する法律、労働組合法施行令（昭和二十四年政令第二百三十一号）、労働関係調整法施行令（昭和二十一年勅令第四百七十八号）及び行政執行法人の労働関係に関する法律施行令（昭和三十一年政令第二百四十九号）をいう。

二～八 （略）

上記では以下のような略語定義がされているものの、文章を解析し、各語句に対応するものを把握する必要がある。

- 「労組法」 → 労働組合法
- 「労調法」 → 労働関係調整法
- 「行労法」 → 行政執行法人の労働関係に関する法律
- 「地方公労法」 → 地方公営企業等の労働関係に関する法律
- 「労組法施行令」 → 労働組合法施行令（昭和二十四年政令第二百三十一号）
- 「労調法施行令」 → 労働関係調整法施行令（昭和二十一年勅令第四百七十八号）
- 「行労法施行令」 → 行政執行法人の労働関係に関する法律施行令（昭和三十一年政令第二百四十九号）

このようなパターンは、難易度は高いものの、機械的な処理での対応を検討する余地もあるため今回の調査ではパターンB（多少複雑な自然言語処理が必要となるもの）として分類している。

定義条文等で定義されるがその後再定義されるもの

商法（明治三十二年法律第四十八号）第四条・第十一条

（定義）

第四条 この法律において「商人」とは、自己の名をもって商行為をすることを業とする者をいう。

2 （略）

（商号の選定）

第十一条 商人（会社及び外国会社を除く。以下この編において同じ。）は、その氏、氏名その他の名称をもってその商号とすることができる。

上記のパターンでは第四条で「商人」が定義されているものの、第十一条でその定義内容が変更され、かつ有効な範囲が限定されている。このようなパターンでは定義語・定義文を抽出に対応できたとしても、その有効範囲や条件を把握して、それに対応した範囲でのリンク情報付与を実施しなければならない。

このため、複雑な機械的な処理の実施もしくは人手による情報の付与などの必要となり、パターンC（複雑な自然言語処理や法制執務的な観点からのメタ情報が必要となるもの）に分類した。

先に定義語が出現しその後定義文が現れるもの

少年警察活動規則（平成十四年国家公安委員会規則第二十号）第六条・第七条第一項

（早期発見）

第六条 第二条第六号から第十号までに掲げる少年については、街頭補導（次条第一項に規定する街頭補導をいう。）及び少年相談を適切に実施し、並びに警察の各部門間及び警察と関係機関の連携を図り、これらを早期に発見するように努めるものとする。

（街頭補導）

第七条 街頭補導（道路その他の公共の場所、駅その他の多数の客の来集する施設又は風俗営業の営業所その他の少年の非行が行われやすい場所において、前条に規定する少年を発見し、必要に応じその場で、これらに第十三条第一項、第十四条第一項、第三十六条第一項、第三十八条第一項又は第三十九条第一項に規定する措置を執る活動をいう。以下同じ。）は、自らの身分を明らかにし、その他相手方の権利を不当に害することのないよう注意して行うものとする。

このパターンでは第六条に「街頭補導（次条第一項に規定する街頭補導をいう。）」とあり、次条（第七条）で規定内容が提示されている。

このように条文を超えた形で定義語と定義文が存在するものも見受けられたが、条文（文章）をまたいだ形で語句の関連付けを把握しなければならないため、本調査研究においては機械的な処理のみでの対応が困難であると判断し、パターンCと分類した。

定義語・定義文よりも先立って条件が来るもの

協同組合による金融事業に関する法律（昭和二十四年法律第百八十三号）第四条第一項

（信用協同組合等の子会社の定義）

第四条 この法律（前条を除く。）において「子会社」とは、信用協同組合等がその総株主等の議決権（総株主又は総出資者の議決権（株式会社にあつては、株主総会において決議をすることができる事項の全部につき議決権を行使することができない株式についての議決権を除き、会社法第八百七十九条第三項（特別清算事件の管轄）の規定により議決権を有するものとみなされる株式についての議決権を含む。以下同じ。）をいう。以下同じ。）の百分の五十を超える議決権を保有する会社をいう。この場合において、信用協同組合等及びその一若しくは二以上の子会社又は当該信用協同組合等の一若しくは二以上の子会社がその総株主等の議決権の百分の五十を超える議決権を保有する他の会社は、当該信用協同組合等の子会社とみなす。

このパターンでは『この法律（前条を除く。）において「子会社」とは、』とあるように定義語・定義文よりも先に『この法律（前条を除く。）において』という有効範囲・条件が出現している。こういった表現の場合は、表記的な観点ではなく内容解釈による部分が大きく、機械的な処理を行う場合、文字列処理や自然言語処理のような範疇を超えた処理の可能性を考慮する必要がある。

他の法令での定義を使用するもの

特定計量器検定検査規則（平成五年通商産業省令第七十号）第一条

（用語）

第一条 この省令において使用する用語は、この省令に特段の定めのない限り、計量法（平成四年法律第五十一号。以下「法」という。）及び計量法関係政令において使用する用語の例による。

このパターンにおいては対象となる法令中に定義語・定義文が存在せず、他の法令を参照するような形を取っている。自然言語処理による定義語・略称の抽出・リンク情報付与といった観点からの調査であるが、こういったパターンの存在を認識することと、目的にあるような規定の対応関係を把握する業務負荷という点からも何らかの機械的な処理で対応が可能であろうという点を考慮しパターンCとして分類した。

5. 4. 内容解釈に起因する課題の整理

5. 4. 1. 課題の類型一覧

内容解釈に起因する課題については、現状、機械的な処理での対応が困難であると考えられるパターンである。

- 括弧書きが入れ子構造や並列配置などで表現されているもの
- 「文脈的な除外」と「法解釈上の除外」の判断が必要なもの

なお『「文脈的な除外」と「法解釈上の除外」の判断が必要なもの』については、具体的な事例は発見できなかったものの、「～条は除く」のような「出現場所のような文脈的な除外」と捉えられるものと「～する者に限る」や「～の場合は除く」のように「法解釈上での除外」と捉えられるものを自然言語処理だけで判断することは困難であると考えられるため、内容解釈を必要とするパターンD（自然言語処理の範疇での処理が難しいもの）と分類した。

5. 4. 2. 課題の類型に関する事例

括弧書きが入れ子構造や並列配置などで表現されているもの

地方公共団体の手数料の標準に関する政令に規定する総務省令で定める金額等を定める省令（平成十二年自治省令第五号）第二条

第二条 令本則の表十七の項の2の下欄の総務省令で定める場合は、次の各号に掲げる屋外タンク貯蔵所の区分に応じ、当該各号に定める場合とする。

一 特定屋外タンク貯蔵所及び準特定屋外タンク貯蔵所（次号に掲げるものを除く。）

屋外貯蔵タンクのタンク本体並びに基礎及び地盤（**地中タンク（規則第四条第三項第四号に規定する地中タンクをいう。）に係る特定屋外タンク貯蔵所及び準特定屋外タンク貯蔵所**にあつてはタンク本体及び地盤、海上タンク（規則第三条第二項第一号に規定する海上タンクをいう。）に係る**特定屋外タンク貯蔵所及び準特定屋外タンク貯蔵所**にあ

ってはタンク本体及び定置設備（規則第四条第三項第六号の二に規定する定置設備をいう。）（定置設備の地盤を含む。）の変更以外の変更に係る消防法（昭和三十二年法律第百八十六号）第十一条第一項後段の規定に基づく変更の許可の申請（以下この条において「変更許可申請」という。）に係る審査の場合
二～五 （略）

上記のパターンの太字部分においては括弧書きによる略称の規定が入れ子構造になっており、かつ条件の指定も添えられている形である。

地方税法第五十三条

（法人の道府県民税の申告納付）

第五十三条 （略）

2～4 （略）

5 第三項の法人を合併法人（合併により被合併法人（合併によりその有する資産及び負債の移転を行つた法人をいう。以下この条において同じ。）から資産及び負債の移転を受けた法人をいう。以下この条において同じ。）とする適格合併が行われた場合又は当該法人との間に法人税法第二条第十二号の七の六に規定する完全支配関係（以下この条において「完全支配関係」という。）（当該法人による完全支配関係又は同号に規定する相互の関係（以下この条において「相互の関係」という。）に限る。）がある他の法人で…

以下略

上記の太字部分のパターンにおいても括弧書きの入れ子構造や並列配置が見られる。

自然言語処理においては、文章の最小構成要素として語句を対象とした係り受け分析を行うことはできるが、法令表記パターンで見られる括弧書き単位での係り受け分析を行うことは現状困難である。

5. 5. 調査結果に基づく考察

既存法令条文の分析と「以下リスト UI」生成用データの作成と構築作業により、以下のような状況を把握することができた。

法令中に存在する定義語・略称の抽出

規定のみの表現であるものについては多くの場合パターンマッチで対応できることが把握できた。しかしながら、有効範囲や条件を伴うものや、表記的な解釈以外の対応が必要なものについては、今後、自然言語処理以外でのアプローチも検討する必要があることも明らかとなった。

定義語・略称の使用箇所の特定とリンク情報付与

前述のような有効範囲や条件、法令中で再定義されるもの等、一意的なルールによる機械的処理で該当する箇所を特定することが困難なパターンが存在することが見えた。

また構築した「以下リスト UI」について、使用箇所全ての定義語・略称にリンク情報を付与したものの、使い勝手という点から「情報が多すぎて見づらく、リンク情報を付与する場合はそのユーザー・インターフェースも検討が必要ではないか」というような気づきもあった。

法令条文特有の表記ルールの存在

表記的な解釈を範疇とする自然言語処理では対応が難しい、法令条文特有の括弧書きの入れ子構造や並列表現が存在することが確認できた。

6. 本調査研究のまとめ及び今後の提案

6. 1. 法令中の併合的・選択的接続詞の表記への自然言語処理のア

プローチについて

ここまでの調査・実証から、法令条文における表記ルールはある程度パターン化できるものの、その性質上内容の解釈が必要なものも多く、自然言語処理の導入のみで対応できるものは現時点では限定的であることが見えた。

既存条文を自然言語処理による解析のみで、「接続詞を含んだ複雑な条文」を構造的に解釈することは、自然言語処理における技術的な点と法令表記ルールの複雑さという点においてまだまだ困難なものであった。

しかしながら、読点の付与ルールなどの法制執務に関連する表記的なルールをある程度包括した仕様の実装、それに加え形態素解析や品詞付与など専用ソフトウェア実行環境によって対応可能な自然言語処理で対応できる範囲の検討を踏まえることによって、今回検証した「条文生成を行う UI」を段階的に実装しながら、法制執務においてどのような情報提示の方法や機能が品質向上・効率化において有用かを検証していくことが、自然言語処理的なアプローチを取り入れた「併合的・選択的接続詞を含んだ複雑な条文」の効率的かつ高品質な作成・解釈へ発展できると考えられるであろう。

6. 2. 定義語・略称に対するリンク情報付与への自然言語処理のアプローチについて

本レポートで扱ってきた、いわゆる「以下リスト」については、これまでも文字列処理によるパターンマッチングなどの機械処理と手作業によるデータ調整で対応してきた背景があった。しかしながら、その性質上、新規の箇所についての都度の対応やその網羅性において課題が生じていた。

今回は文字列処理によるパターンマッチングと自然言語処理を含めた機械処理での対応、つまり「手作業によるデータ調整」を省いた形で処理できる範囲を調査したものの、現時点では一部のパターンを除き、様々なパターンを網羅したものを提示することは困難であることが分かった。

この点において、機械的な処理のみで逐次的な形で情報提示ができる仕組みづくりにいきなり取り組むのではなく、ある程度の準備期間を経てデータを構築するという方向性を前提に、自然言語処理を活かした「語の定義箇所の抽出」、「定義語・略語と思われる箇所の抽出」をサポートできるようなツールを活用した手作業によるデータ構築から取り組むのが現実的な対応と考えられる。

また、すでに情報が付与された既存データがあれば、それらを解析し、扱うデータへ変換し、法制執務に活用できる形にすることもひとつの手段ではある。このときにも自然言語処理的なアプローチが有用であると考えられる。

さらに今回検討してきたような「精度の高い情報提示」とは別に「リンク情報が付与される箇所の多さ・範囲」などの法制執務に関わるユーザーの使いやすさに直結する点も十分に検討する必要がある。

6. 3. まとめ

これまで述べてきたように、自然言語処理を活用し、法制執務の業務フローの効率化・品質向上を目指す場合、自然言語処理に関する技術的な限界（特に辞書やコーパスの充実というところによることが大きい）や表記ルールのより広範にわたる整理分析に依存するところが非常に大きいことが明確になった。

そのため現時点では「自然言語処理を活用したシステムを開発して業務効率化を図る」よりも「自然言語処理を活用したデータ構築を行い、そのデータを利用して業務効率化を図る」といった形のほうがより現実的な取組となることも整理できた。具体的には「法制執務における内容解釈を補助するツール」のような形で自然言語処理によるメリットを活かすような取組が想定される。

その後、取組を更に発展させる方向性としては、構築されたデータを自然言語処理やAIなどの強化学習に取り入れることで「法制執務における内容解釈を補助するツール」の性能をさらに向上させ、また、併せて法令表記ルールの見直しを進めることで、「ツールからデータへ」、「データからツールへ」といった再帰的な開発アプローチで法制執務における自然言語処理の活用を進めていくことも考えられる。

そしてその先には、こういった成果による「自然言語処理を活用したシステムを開発して業務効率化を図る」道筋が見えてくるだけではなく、本レポート中でも述べた「文章のオブジェクト表現」を実現し、法制執務の品質向上・効率化を超え、さらに社会全体の価値創出につながる「Rules As Code¹⁴」のようなより発展的な法令条文の在り方へ繋がられるはずである。

¹⁴ 機械実行可能な形式で法令などのルールを記述し、法令の効果を事前にシミュレートしたり、法令を直接システム実装したりすることで、政策立案や行政運営の効率化・合理化を目指す手法や理念のこと