

令和 6 年 5 月 23 日現在

機関番号：38001

研究種目：若手研究

研究期間：2020～2023

課題番号：20K18854

研究課題名(和文) 臨床研究を対象とした複数の法令間の特徴と関連性の抽出とそれらの可視化

研究課題名(英文) Extraction and visualization of relation and characteristics for multiple laws regarding clinical research

研究代表者

大城 絢子 (Ohshiro, Ayako)

沖縄国際大学・経済学部・准教授

研究者番号：70773399

交付決定額(研究期間全体)：(直接経費) 700,000円

研究成果の概要(和文)：本研究は、臨床研究関連の複数の法令の可読性向上のために、法令間の特徴と関連性を抽出および可視化することを目的としている。対象の省令や法、指針に対し1. 単語一致度を用いた各法令間の条レベルでの類似性評価 2. 単語一致度の其他法令への適用 3. 「自己相互情報量と単語群一致度を組み合わせた手法」による臨床研究関連法規の類似性評価の検討 4. 3の手法の新聞記事への適用 5. 共起ネットワークを用いた臨床研究関連法規の可視化に取り組んだ。
研究成果として、対象法令の可視化・対象法令およびその他の法令間の類似性評価・類似性評価手法の提案・テキストデータの特徴に応じた類似性評価手法の選定の可能性を示した。

研究成果の学術的意義や社会的意義

データの特徴に応じた言語処理手法の選択の可能性が示された。また大学等の研究教育機関や地域医療機関において臨床研究を初めて実施・計画または支援する医療者だけでなく患者・国民にとって、本報告が法理解の助けになれば幸いである。

研究成果の概要(英文)：The purpose of this research is the extraction and visualization of relationships and characteristics from multiple laws regarding clinical research in order to improve readability. For the target ministerial ordinances, laws, and guidelines, the following tasks were undertaken: Task1. Assessing the article-level similarity between each law using word matching. Task2. Extending the application of word matching to other laws and regulations. Task3. Propose a method by the combination of Pointwise Mutual Information and Word Pair Matching. Task4. Adapting the method from task 3 to analyze newspaper articles. Task5. Creating a visualization of clinical research-related laws and regulations using a co-occurrence network.
We demonstrated the possibility of visualizing the target law, evaluating the similarity between the target law and other laws, proposing a similarity evaluation method, and selecting a similarity evaluation method according to the characteristics of text data.

研究分野：知能情報工学

キーワード：臨床研究 法情報 言語処理

1. 研究開始当初の背景

妥当性のある医療の実施や医療の質の向上には、診療上の疑問を解決する臨床研究が必修である。1997年には「医薬品の臨床試験の実施の基準に関する省令(Good Clinical Practice)」、2017年の「人を対象とする医学系研究に関する倫理指針」改訂に加え2018年には「臨床研究法」が施行され、臨床研究はこれらの省令・指針・法に沿った研究計画と遂行が求められるようになった。特に臨床研究法の施行直後には、すでに施行されている省令や指針に加えて法への対応も求められ、臨床研究関連の法規を、各ポイントに沿って整理し解説された論文が多く作成され、これらの関係を端的に表現することの必要性が窺われた。一部の大学病院では臨床研究の法務担当を配属し、大学および地域医療機関、さらには企業で実施される臨床研究が法基準を満たすかのチェック、研究者を対象とした説明会などをおこなうなどの対応をした。しかし法務担当者の負担や法務管理コスト軽減のためにも、研究者である医療者自身が法を理解していることが理想である。

臨床研究に関連した各法規やこれらの関係を何らかの形で簡素化および可視化することで、臨床研究に携わる医療者や実務者にとって法理解の助けになり、また法務管理コストの軽減も期待できる。

2. 研究の目的

臨床研究を対象に施行された複数の法令の可読性向上のために、法令間の特徴と関連性を抽出し可視化する。

3. 研究の方法

本研究では、自然言語処理手法を該当の法令データに適用することで各法令の特徴および法令間の関係を可視化する。また法令間の可視化のために用いる類似性評価尺度を提案し、さらにその評価尺度を、該当法令以外の法令データや法令以外のテキストデータへ適用し有効性を評価する。その方法について下記に述べる。

A. 各法令の特徴抽出

A-1. 各法令文のロジック構造の抽出

A-2. 共起ネットワークを用いた各法令の可視化

B. 各法令間の関係抽出

B-1. 共通形態素や文字列類似度を用いた法令間の関連性の可視化

B-2. 法令間の参照関係の抽出

B-3. 単語一致度を用いた各法令間の条レベルでの類似関係の計算

B-4. 自己相互情報量と単語群一致度を組み合わせた法令間の類似性評価手法の検討

C. 臨床研究関連法令以外の法令およびテキストデータへの適用

C-1. 単語一致度のその他の法令への適用

C-2. 「自己相互情報量と単語群一致度を組み合わせた類似性評価尺度」の、法令以外の文書データへの適用

4. 研究成果

A. 各法令の特徴抽出

A-1. 各法令文のロジック構造の抽出と法令間の類似度計算

法令文はある一定のルールに従って矛盾なく設計され規定されているためロジック構造での表現が可能である。ロジック構造を導入し法令特有の表現を除去するという流れを標準化することで、本研究の対象法令に限らず一般的な法令データを扱う場合も、ある程度意味的精度の保証された関連性が示すことができ本来の法令間の類似関係を求めやすくなることが期待できる。ロジック表現した法令間の類似関係を条文間ごとに算出した。

A-2. 共起ネットワークを用いた臨床研究関連法規の可視化

共起ネットワークを用いて、臨床研究を対象とした省令・法令・指針のそれぞれについて、含まれる単語間の関係を可視化した。(図 1) 特に強い共起関係にある単語の組み合わせ・関連法規における共通出現単語やそれらの各ネットワーク内における共起単語を比較し、各法規を単語レベルで予測することで、従来の記述的アプローチとは異なる視点で臨床研究関連の法規を解釈できることが示された。他方で、本手法を用いて省令・法令・指針を完全に理解することは困難であり、各法規レベルでの分析が必要であることも示された。

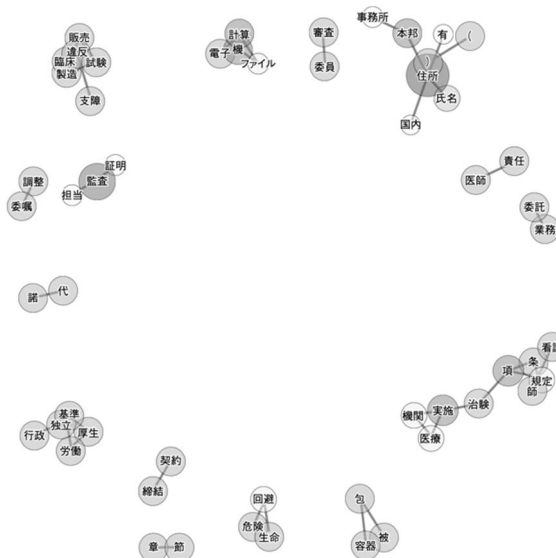


図 1. 「医薬品の臨床試験の実施の基準に関する省令 (Good Clinical Practice)」の共起ネットワーク

B. 各法令間の関係抽出

B-1. 共通形態素や文字列類似度を用いた法令間の関連性の可視化

各法令文に含まれる形態素(表層語)レベルに落とし込み、図 2 のように共通形態素を抽出することで法令間の関連性を可視化した。

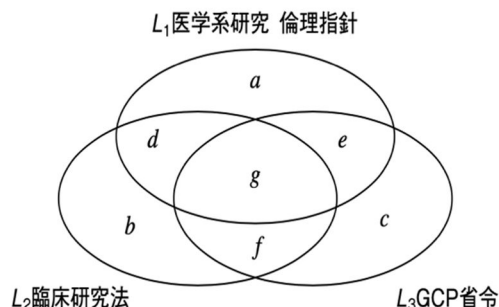


図 2. 臨床研究に関連した 3 法令の形態素の重複状況

また文字列間の一致度を文字列類似度とすることで、形態素の完全一致レベルでは抽出できなかった「類似性」を求めた。

B-2. 法令間の参照関係の抽出

法令内において同一表現の多用を避けるために条文の参照関係を整理するプログラムを実装した。

B-3. 単語一致度を用いた各法令間の条レベルでの類似関係の計算

共通単語数が多い文章は類似度が高いという仮説のもと、「単語一致度」を用いて臨床研究に関連した 3 法令の各条文間の類似度を予測した。その結果、意味的に「類似性が高い/一致」である組み合わせが正しく獲得された一方で、「類似性が低い/不一致」である組み合わせの条文の類似度が高いとされたケースもあることがわかった。

B-4. 自己相互情報量と単語群一致度を組み合わせた臨床研究関連法規の類似性評価手法の検討

自己相互情報量・単語群一致度・単語群発生頻度を組み合わせた類似尺度を提案し、臨床研究を対象とした 3 つの法令間の条文の類似性を数値化した。性能比較実験の結果、既存の類似尺度より類似度推定精度が改善された組み合わせがあること・既存の尺度との組み合わせによりさらに改善が期待できることを確認した。

C. 臨床研究関連法令以外の法令およびテキストデータへの適用

C-1. 単語一致度のその他の法令への適用

沖縄県の本土復帰にあたり施行された地域振興法と、沖縄県以外の地域に施行された

地域振興法を比較するために、B-3 で述べた「単語一致度」を用いてこれらの関係を定量化し可視化した。単語の重複率をもとに各地域間の振興法の特徴を比較した。

C-2. 「自己相互情報量と単語群一致度を組み合わせた類似性評価尺度」の、法令以外の文書データへの適用

B-4 で示した「自己相互情報量と単語群一致度を組み合わせた類似性評価尺度」を新聞記事データに適用し、提案尺度と従来手法による尺度を用いて、新聞の各記事間の類似性を評価した。従来手法には Deep Learning による手法、集合理論や情報理論 を取り入れた手法を用いた。性能比較対象として臨床研究関連の法令を選択した。扱うテキストデータのジャンルによって精度が異なることがわかった。解析に用いるテキストデータのジャンルによって適用する手法を分けることで、より良い類似性の評価ができる可能性が示された。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 大城 絢子, 岡崎 威生, 植田 真一郎	4. 巻 55
2. 論文標題 共起ネットワークを用いた臨床研究関連法規の可視化	5. 発行年 2024年
3. 雑誌名 臨床薬理	6. 最初と最後の頁 57, 62
掲載論文のDOI (デジタルオブジェクト識別子) 10.3999/jscpt.55.1_57	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 大城 絢子, 岡崎 威生, 植田 真一郎	4. 巻 54
2. 論文標題 単語一致度を用いた臨床研究関連法規の関係可視化の検討	5. 発行年 2023年
3. 雑誌名 臨床薬理	6. 最初と最後の頁 43, 48
掲載論文のDOI (デジタルオブジェクト識別子) 10.3999/jscpt.54.1_43	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 大城 絢子, 岡崎 威生, 植田 真一郎
2. 発表標題 自己相互情報量と単語群一致度を組み合わせた臨床研究関連法規の類似性評価の検討
3. 学会等名 言語処理学会 第29回年次大会(NLP2023)
4. 発表年 2023年

〔図書〕 計1件

1. 著者名 宮城和宏, 名嘉座元一, 比嘉正茂, 鹿毛理恵, 安藤由美, 高哲央, 大城絢子, 照屋翔大, 吉原千鶴, 小濱武	4. 発行年 2023年
2. 出版社 編集工房東洋企画	5. 総ページ数 329
3. 書名 「復帰」50年と沖縄経済(p.199-p.223 データサイエンスで見た復帰50年)	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------