

# 地方議会議事録の統合と利活用に関する社会情報学的研究（継続）

代表研究者 渡部 春佳 NIRA 総合研究開発機構 研究コーディネーター・研究員  
共同研究者 浅谷 公威 東京大学 工学研究科 特任講師

## 1 はじめに

2023年4月統一地方選において、市町村議選・町村長選の投票率は過去最低を記録した。地方議会の存在意義が問い直される中、これまでの議論の蓄積から、地方自治や政策形成にとっての地方議会の意義を検証していくことは重要である。二元代表制に基づく地方自治において地方議会では、地域住民から選出された「首長」と「議員」の間の相互作用による意思決定が行われる。また、議会は政策形成や選択に一定の影響を持ち、執行機関である自治体が市民に対するアカウンタビリティを果たす重要な場でもある。昨今、社会科学・政治学の分野では、機械判読可能なデータの増加に対する人工知能の活用が期待されている。政治学の分野では「Text as Data」といい、政治や政策に関するテキストを分析データとして利用することにより、潜在的な政策課題（政治的争点）の探索や、政治家のイデオロギー配置を検証する研究が行われてきた（Grimmer and Stewart 2013）。本稿では、国内において公開やデータベース整備が進みつつある地方議会議事録に焦点を当て、政策研究のための統合と利活用のための方向性を検討する。

日本において、地方議会に関する情報公開は2001年4月施行された「情報公開法」以降、進められてきた（大山 2007）。議事録の公開によって、地方議員活動の監視・可視化ができるほか、地域の政策課題・争点を具体化することによる、市民への資源共有という意義がある（土山 2018）。実際に、地方議会議事録を利用した研究として、地域における政策課題の抽出や、議会の存在意義の実証的検証など多方面での活用がなされている。膨大な議会議事録データについては、発展の著しい自然言語処理の手法を用いた分析が期待される。

本稿では、自治体横断的な議事録データの収集とテキスト分析を実施する。研究手順として、分析手法を検討するため、量的手法を取り入れた議事録データを用いた研究動向について、その目的・分析手法・分析対象の観点からまとめる。そして、事例研究として特定の劇場に言及された発言を対象として、文化政策という政策領域においてどのような議題が含まれるかトピックモデルにより明らかにする。以上の既存研究の調査と事例研究の結果を受け、公開されている議事録データの特性を踏まえ、その横断的な収集・分析によって現状でどのような研究が可能で、今後、社会科学・政治学研究として一定の知見を導き出すうえでの課題は何かについて議論する。本稿の構成は以下の通りである。まず、第2章で地方議会議事録の政策研究への利活用に関する課題を概観する。第3章では、地方議会議事録を扱った既存研究の調査からこれまでどのような狙いのもとで研究がなされ、どのような方法・対象が用いられたかを示す。第4章では、事例研究として本研究において収集された47都道府県議会議事録データに対して、文化施設・劇場に関する発言に絞った分析結果を示す。第5章で、2年間の研究を踏まえ、地方議会議事録の統合と利活用に向けた研究の課題や可能性について議論する。

## 2 地方議会議事録の政策研究への利活用に関する課題

現在、地方議会での発言を収めた議事録は、各自治体のHPで議事録検索システムから閲覧することができる。小樽商科大学木村泰知らによって「地方議会会議録コーパスの構築とその学際的応用研究」が実施され、2022年現在は、2011年4月～2019年3月を対象とした都道府県議会（および2015年～2019年の東京23区議会）の議事録を検索できるようになっている（<http://local-politics.jp/>）。このほかに、議事録だけでなく、審議の様子を収めた映像をHP上で公開している自治体もある。このように、昨今では議会議事録の公開や自治体横断的な検索システムの整備が進められている。しかし、増田（2016）は、「最近の地方議会審議の急速な『見える化』にもかかわらず」、「映像やデータを十分に活用できていないことは明白」と指摘する。

活用が進みづらい背景の一つには、第3章でも言及するように、各自治体に公開された議会議事録データが、分析データとして統一的な構造になっていないことがある。例えば高丸（2019）は、発言者情報の付与

が不足していた問題などがあるとし、発言者の特定や名寄せなどの必要性を指摘する。また、発言者情報付与の成果をいかして、議員発言の特徴語の抽出が行われている。もう一つに、膨大なテキストデータの要約という問題がある。議事録システムから得られる情報は、誰がいつどのような発言をしたかということである。何十年にもわたる膨大なテキストデータを読み解くためには、できる限り人の手による労力をかけず、個々の発言について文脈を理解しながら要約するということが必要となる。

本稿は、特に後者の地方議会における議論の要約に関して、第4章の事例研究においては、特定の政策分野に焦点を当て、議論される議題を自動的に抽出することを試みる。具体的には、限られた事例に対して行われた渡部(2020)を発展させる形で、対象を47都道府県の議会議事録に広げ、それらにトピックモデルを適用し、自治体間で横断的に議論されているトピックを自動的に分類・構造化することを行う。政治情報をトピックモデルにより自動的に抽出する方法論については、要約・解釈が調査者の目的に左右されることが指摘され、複数の側面から結果をみることで妥当性を検証することが必須である(Isoaho et al. 2021)。そのため、本研究でも時系列・地域間でのトピック出現傾向をあわせてみることで分類結果の妥当性について検討し、当該研究潮流へ貢献することを目指す。

### 3 地方議会議事録を用いた研究の動向

本章では、地方議会議事録を用いた既存研究の調査結果を示す。調査方法としては、CiNii, google scholar を併用し、「議会議事録」のキーワード検索で既存研究の収集を行った。日本の地方議会を対象とすることから、国内の研究を対象とした。対象としては、地方議会議事録(本会議・委員会)を対象としているもの(国会、省庁の諮問協議会等は含まない)とし、学術論文、書籍収録論文、紀要論文の他、学会予稿集でも分析内容の把握できるものについては含めた。収集作業は研究補助者1名と筆者が行い、内容確認は筆者が行った。

収集論文件数は51件で、調査内容としては以下のような項目を設定した。まず、①地方議会議事録を扱う狙いについてである。議会議事録を用いた研究には、政治学理論の検証か、特定政策領域の議論であるか、課題探索的/仮説論証的か等様々なものが想定されるが、実際はどうであるかを確認した。次に、③対象とされる会議が複数地域か単一地域かについて、また②対象となる会議の種類について調査した。さらに、④分析手法が、定量/定性/その両方であるか、定量分析の場合は、人力か機械処理を用いているか(テキストマイニング)確認した。そして、⑤対象とした発言(全体、議員や行政職員の答弁等のみであるか)ということを確認し整理した。以下では、これらの概要について、収集した論文のうち代表的な論点を含むと考えられるものに言及しながらまとめる。

まず、①地方議会議事録を扱う狙いについてである。はじめに、地方議員活動の監視・可視化という面がある。選挙公約と議会活動の不一致度の検証(長谷川 2016)や、多選議員のモチベーション低下(内田ら 2019)を指摘した研究がある。次に、政策課題の探索や政策形成過程の研究がある。政策全体もしくは特定政策領域の課題の抽出するもの(増田 2017; 渡部 2020)、類似政策間での重点に対する自治体間比較を行うもの(渡部 2020; 石川・梅村 2021)、政策決定に対するアクターの影響力を把握するもの(廣谷 2021)、会派の政策的同質性を検証するもの(増田・爲我井 2019)がある。その他に、自治体経営に必要な情報(内容・発信経路)の明確化(本田 2011; 2012)や、議会の持つチェック機能に着目した研究(宮坂 2019; 横田 2022)、行政・議員間関係の検証(増田 2018)、議員の判断根拠の明確化(鈴木 2018)、議会改革への態度推定(増田 2016)など様々な研究がみられる。

②対象となる会議としては、単一自治体の地方議会か複数地方議会を扱っており、複数の場合は、特定地域の市町村、一部の都道府県、一部の政令指定都市などを対象としていた。また、③分析の対象とされている議会の種類は、地方議会の会議全体の場合もあれば、提出議案等の影響が少ないという理由で本会議の定例会や臨時会に限定するもの(橋本 2011)、定例会を対象とするもの(田中 2020; 梶原・名取 2020)、決算審査特別委員会のみ対象とするもの(横田 2022)と多様である。また、研究目的によっては、当初から分析対象を特定テーマについて議論されている会議に絞ったり、検索システムから特定のワードにヒットする発言に絞ったりするという例もみられる。

④手法としては、テキスト分析をはじめとする定量的な手法が主である。定性的な手法を採るものであっても、定量的手法を部分的に取り入れているものが大半である。代表的なものとしては、議事録検索システムを用いたキーワードの検索結果の集計を行うもの、発言件数の数え上げを行うものがある。また指標を作

成することで自治体間比較を試みるものとして、松本（2008）は広範度（地域間に渡る質問か）、集中度（同一の自治体で繰り返しされる質問か）、変動度（時期による変動がある質問か）の指標作成を行っている。さらに、議事録から得られた情報を用いた回帰分析するもの（長谷川 2016）もあった。主要な定量的な手法としては、テキストマイニングによるものがみられる。共起ネットワークや階層クラスタリングを行うもの（増田 2017；渡部 2020 ほか）、ネガポジ分析を行うもの（梶原・名取 2020；名取ら 2020）があった。そして、⑤対象とした発言としては、全体、議員の発言、さらに議員の一般質問のみ（松本 2008）、質疑応答部分のみというものもある。分析単位としても、会議ごと・発言ごと・段落ごと等様々である。

最後に、これらの分析を実施する研究のほか、コーパス整備に関する動向をまとめる論考も収集した。例えば、地方議会議事録のデータ収集方法における問題点や公開形式が指摘したもの（井原ら 2017）や、現状の議会議事録の不足に対し、議員情報（氏名・役職・選挙区・生年・性別等の情報）の補足を行う実践（内田ら 2019）がある。同様に、名取ら（2016）はデータテーブルの作成を行い、分析の際に審議内容に加えて変数として扱うことのできる要素を整理し、比較政治学的な関心のもとでの研究を可能にするデータベースを構築している。

以上、国内において、公開やデータベース整備が進みつつある地方議会議事録に焦点を当て、議事録データを用いた研究動向について概観した。地方議会議事録を扱う狙いは多岐にわたり、その目的によって分析手法も異なっていることが明らかとなった。総じて、地方議会を「行政・議員間の議論による政策選択・決定の場」「行政活動のチェックの場」「議員の活動の場」のいずれの面に着目するかにより分析対象や手法は異なってくる。そして、これまでの調査の中で指摘されている課題としては、分析から導き出された傾向を一般化できるほどのサンプルを対象としていなかったことが指摘されてきた。現在の言語処理技術を用いてすれば大規模な議事録データを分析することはできるが、発言者の特性や議会の構成状況を踏まえ、精緻な分析を行おうとすればするほど、分析する議事録データベースの整形に労力がかかる。また、語の出現頻度に基づく分析では、会議ごとの特性に左右される点が挙げられている。これについては、発言の文脈や語の辞書的意味を踏まえることのできる言語処理技術によって一定補完することが可能だろう。実際に、従来は人力によるカテゴリズや KHcoder による一定規模のテキストデータを対象とした分析が主であったのに対して、概念体系の自動構築の可能性を指摘する研究（長谷川ら 2008）や、TF・IDF によって求めた都道府県別の特徴語の抽出（高丸ら 2017）などの試みが行われている。最近では、自然言語処理的手法を取り入れた、より大規模な議会議事録のデータ分析に適応しうる試みもある。具体的には、政治課題の主辞（促進、制定、改正など政治課題を発言する際に意味的に中心になる単語）となる名詞候補を設定することで、複雑に絡み合った政治課題を自動的に特定する調査が行われ、7 割の精度で政治課題を抽出したとしている（木村 2021）。今後の活用に向けた課題としては、自治体横断的なデータベースの構築とともに、全体的な議論の要約・抽出方法の提案が待たれているといえる。そこで本稿は後者にあたる、言語処理技術を応用した議論の要約に取り組む。

## 4 「47 都道府県議会議事録」を用いた事例研究

本章は、47 都道府県の議会議事録の議論要約を通じた考察を行っていく。まず、用いるデータと研究方法の説明を 4-1 で行い、4-2 で実際の議会で議論に上がっているトピックの抽出と、その時系列推移の分析を実施する。

### 4-1 研究方法とデータ

分析データとしては、都道府県議会議事録データの横断的に収集したものに基づく。具体的には、収集時点（2022 年 2 月）までに公開されている 47 都道府県の地方議会議事録（本会議、委員会、最も古いもので 1947 年～）の発言を収集した。収集の単位は「段落」である。収集したデータからは、都道府県名、会議名、会議開催年月日、発言、発言者からなるデータベースを作成した。

研究方法として、議会議事録に対するテキスト分析を行う。まず、分析データとして、2021 年 2 月までに公開されている 47 都道府県の地方議会議事録の発言とその関連情報（会議名、会議開催年月日、発言者）を収集した。実際の議事録では、特定の発言者による 1 回の発言であっても、複数の議題に言及し、長くなる場合がある。本研究は特定政策課題の分析に適した形でのデータを収集できるよう、収集単位は段落とした。実際に収録されている会議については、1 年目の研究成果報告書の渡部・浅谷（2022）を参照され

たい。

さらに収集された議会議事録に対して、文化施設・劇場という特定政策領域に関するデータの抽出を行った。図1のような政策テーマに関連する単語を含む段落を特定し、本稿はこのうち文化施設・劇場に関連する語のいずれかを含む約32千段落を対象に分析を行った。一定の議論を含む発言（71字以上999字以下）に対し分析を行い、30のトピックと各トピックの特徴を示す代表語を抽出した。



図1 文化施設・劇場に関する発言の出現状況

トピックの抽出にはトピックモデルを用い、その詳細は以下の通りである。LDAという文書の集合を記述するトピックのセットを抽出するための統計的学習技術 (Bleiland et al. 2003) を用いる。さらに、Concept-topic approach (Steyvers and et al. 2010) というトピックモデルに人間定義の概念セットを組み合わせたアプローチを採用する。具体的には、Top2vec (Angelov 2020) を用いた。Doc2Vec もしくは Bert 等によって事前に学習した文書ベクトルを次元削減し、トピックをクラスタリングする。これはヒューリスティックな手法であり、LDAなどに比べて分類精度やトピックの独立性が高い利点がある。

#### 4-2 トピックの抽出と時系列推移

まず時系列・地域別にみた発言を量的に概観する。図2は、時系列でみた発言数の推移であり、左が文化施設・劇場に関する語を含む段落の出現状況、右が収集議事録全体での推移である。文化施設・劇場に関する発言は2005年や2015年前後に出現が増えている。

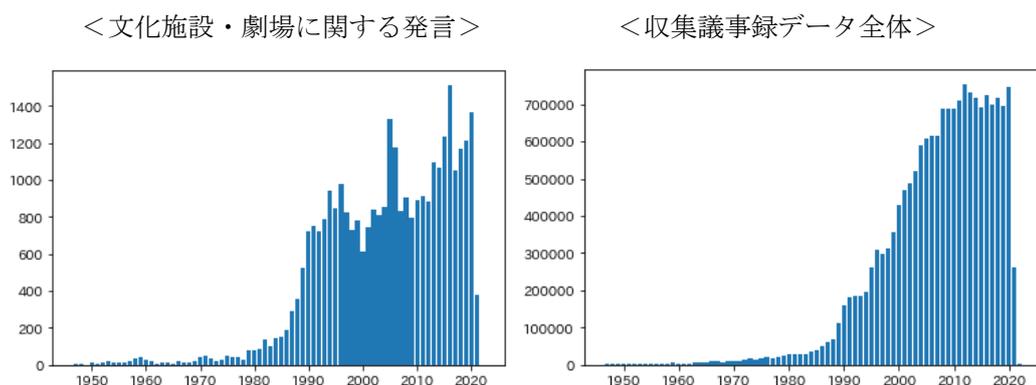


図2 時系列でみた発言数の推移 (単位: 段落)

さらに地域別での発言数を比較したところ東京都が文化施設・劇場に関する発言が4,705段落と、収集議事録全体と共通して圧倒的に多くみられていたが、それ以下では文化施設・劇場に関する発言は高知県、石

川県に多く、これは収集議事録全体の傾向とも異なっていた。このように、発言量からは政策領域特有の結果がみられていることがわかる。

続いて、収集された発言においてどのような議題が現れているかをみるために、約3万段落の発言を全データに対し、形態素解析（Ginza）を行ったうえで、一定の議題が展開されている段落（71字以上999字以下）に対しトピックモデルを適用し、文章の類似度から発言を約1万トピックに分類した。そこから意味的に解釈可能なトピックに落とし込むために、トピックを30件に統合・検討した。

図3は、文化施設・劇場に関する発言に関してトピック分類をした結果である。図中の単語はそれぞれのトピックに関する特徴語であり、冒頭にある数字を各トピックを指名する際の番号とする。これを見るとまちづくりや魅力、にぎわいに関する語が上部のトピック22, 9、指定管理者や管理運営に関する語がさらにその下のトピック1、建物の改修や老朽化に関するものが中央のトピック25、文化芸術活動や芸術ジャンルに関する語が下部のトピック2, 7, 0などにみられている。



図3 文化施設・劇場に関する語のトピック分類結果

例えば、図4はトピック24（左図）とトピック1（右図）のワードクラウドである。この図では、それぞれのトピックに含まれる発言に出現する頻度が他のトピックに含まれる発言に比べ高く、そのトピックに代表的といえる語（特徴語）が大きく表示されている。このような作業を経て、各トピックの示すテーマのタイトルを振った。タイトルについては後述する表1に示した。



図4 各トピックの特徴語のワードクラウド

続いて、これらのトピックが時系列でみてどう出現するかを確認する。表1は、トピック毎のテーマと時系列的にみた各トピックの出現頻度である。表のトピックの順序は、図3の順序に対応している。

表 1 トピックの時系列でみた出現の推移

トピック名		出現の推移 (1980年以降)	頻出年
トピック21	高齢化・環境		1992年
トピック20	(地域) 社会参加		1996年
トピック18	学校		2002年
トピック29	ホール・美術館		2016年
トピック8	コンベンション		2016年
トピック4	土地利用		2006年
トピック22	空間・環境		2009年
トピック19	回遊性・にぎわい		2016年
トピック17	交通		1994年
トピック9	都市 (機能)		1992年
トピック24	使用料・観覧料		2020年
トピック1	指定管理者制度		2005年
トピック28	ホール・感染		2020年
トピック25	改修・老朽化		2005年
トピック16	耐震		2005年
トピック10	問題		1993年
トピック3	問題(2)		2005年
トピック26	経費		2016年
トピック13	財政		1996年
トピック27	---		2005年
トピック11	情報発信・誘客		2016年
トピック6	公演開催		2016年
トピック23	展示開催		2005年
トピック15	展示開催(2)		2017年
トピック14	博物館・劇場・美術館		2006年
トピック2	文化活動・地域文化 (振興・継承)		2015年
トピック7	舞台芸術ジャンル		2003年
トピック0	ホール・美術館		2016年
トピック12	芸術		1994年
トピック5	ホール		2016年

表の「出現の推移」の列に出現推移を図示した。各トピックに含まれる発言数は総量が異なるので、トピックごとに時系列で見た出現の多寡を調べた。青い折れ線は、議事録公開が一定確認できる 1980 年以降の各トピックの出現の推移を示し、赤い点はトピックの出現頻度が最も高かった年を示している。赤い点が実際に指す年は、最右の列に記載している。これを見ると、高齢化・環境（トピック 21）社会参加（トピック 20）は、1990 年代に頻出した。指定管理者制度（トピック 1）、改修・老朽化（トピック 28）、耐震（トピック 16）は 2000 年代に頻出した。そして、コンベンション（トピック 8）、回遊性・にぎわい（トピック 19）、使用料・観覧料（トピック 24）、ホール・感染（トピック 28）、情報発信・誘客（トピック 11）は 2010 年代以降に頻出していた。

本結果から読み取れる論点として、1 点目にまちづくりの文脈において文化施設・劇場が語られる際の焦点の変化がある。1990 年代は住民・社会参加というトピックが出現していたのに対して、2010 年代誘客、にぎわい、アフターコンベンションといったトピックの出現頻度が高まり、語られる文脈が変化していることが予想される。2 点目に、文化施設・劇場に関する議論において、2000 年代に指定管理者制度、施設老朽化への対応など新たな課題の登場していることが指摘できる。さらに、2020 年以降には新型コロナウイルスへの対応という時節性の話題や利用料金についての議論しており、地域内部の課題から、広く外部環境への対応という議題の性質の変化を指摘できる。

本章は、47 都道府県の公開されている地方議会議事録を収集し、それらテキストデータに対する内容分析を行うことで、本研究手法の応用可能性について考察した。本研究の意義として、トピックモデルの利用による、議論の要約にかかる労力の省力化の可能性を示した。特に今回は辞書的意味を踏まえることのでき

る言語処理技術によって、各発言のトピックをヒューリスティックに探索し、多様な議論がなされる各会議での発言を 30 のトピックに要約した。そして、各トピックの時系列推移を見ることで、本研究で行なったトピックの解釈が実際の出来事と大きく乖離していないことを確認した。本手法が各段落における発言をどのようなテーマに関連するかを定量的に明らかにする方法として一定の有効性があることを示すことができた。また、本稿で示した時系列分析以外にも、本研究では都道府県ごとでの各トピックの相関をみることを行い、トピック抽出の妥当性を検討した。

## 5 まとめ

以上、まず第 2 章では地方議会議事録の政策研究への利活用に関する課題を、公開されているデータ構造と発言の要約という 2 点にまとめた。続いて、第 3 章で地方議会議事録を扱った既存研究の調査から、研究目的と研究方法についてまとめた。研究の目的に合わせて、分析対象とする会議や発言を選定していることを指摘した。一方で、研究の目的に合わせた議事録データの取得・整形には、現段階で多大な手間がかかることが見てとれた。そして、第 4 章では事例研究として本研究において収集された 47 都道府県議会議事録データに対して、文化施設・劇場に関する発言に絞った分析結果を示した。その結果、政策領域について、トピックの形で議論を要約することが可能であり、トピックの時系列の出現頻度などから議論の傾向を検討することが明らかになった。

最後に、事例とした政策領域での分析結果から、今後の課題と研究の方向性についてまとめる。一方で、本分析での課題としては、最低限のデータベース整理で分析を行ったゆえの限界がある。全議論を対象としているため、自治体ごとの公開されている会議の種類・時期のばらつきがあり、それが自治体間のトピックを比較する際にも影響している可能性はある。

また、本分析では、自治体が共通に管理する対象として、文化施設・劇場に言及した全発言を対象としたことによって、本対象に関連する全般的なトピックを客観的に明らかにすることができた。今後、本分析結果を政策研究に繋げていくためには、各トピックの多寡が実際に何を示しているのか、すなわち自治体の力をいれている政策課題なのか、政治的争点なのかといった詳細も明らかにすることが必要である。各自自治体の本政策領域に関する予算規模、計画・法整備の進展具合など、外部データを用いることでトピックの解釈を深めていきたいと考えている。また、今回、収集できた会議録について自治体間で量的に差があり、その差は公開されている議事録全体の規模とは必ずしも一致していなかった。これについても、各自自治体におけるトピックの量的規模が、住民の関心をどれだけ反映しているのか、外部データを併用し分析していくことが本研究成果を解釈する上で有用であると考えられる。

## 【参考文献】

- Blei, D., A. Ng, and M. Jordan, 2003, "Latent Dirichlet allocation Journal of Machine Learning Research (3). (<https://www.jmlr.org/papers/v3/blei03a>).
- Dimo Angelov, 2020, "Top2Vec: Distributed Representations of Topics" (<https://doi.org/10.48550/arXiv.2008.09470>).
- 長谷川武三, 2016「地方議員の選挙公約の一致度分析—二元代表制の与野党効果—」『日本地域政策研究』17, 38-45.
- 長谷川大・乙武北斗・木村泰知・洪木英潔・高丸圭一・荒木健治, 2008「市議会会議録を対象とした概念体系構築へ向けた分析」『情報処理学会研究報告』90, 23-28.
- Grimmer, Justin, and Brandon M. Stewart, 2013, "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political analysis*, 21(3): 267-297.
- 橋本武, 2011「議会会議録に見る都道府県の国土計画に対する関心の変化」『計画行政』34 (2): 62-69.
- 本田正美, 2011「地方議会が求める自治体経営情報」『経営情報学会全国研究発表大会要旨集』.
- 本田正美, 2012「地方議会の審議を介した自治体経営情報の明確化 自治体アセットマネジメント導入に関する長崎県大村市議会の事例研究」『経営情報学会 全国研究発表大会要旨集 2012 年春季全国研究発表大会』.

- 廣谷貴明, 2021「政策アリーナ・アクターに着目した教育政策過程の分析—神奈川県横浜市を事例として—」『社会情報研究』3 (1): 1-11.
- Isoaho, Karoliina, Daria Gritsenko and Eetu Mäkelä, Topic Modeling and Text Analysis for Qualitative Policy Research, *Policy Studies Journal*, 49(1): 300-324.
- 石川俊之・梅村輝, 2021「県議会・市議会の会議録からみた「うみのこ」・「やまのこ」・「たんぼのこ」の自治体のかかわり方の違い」『滋賀大学環境総合研究センター研究年報』18(1): 13-20.
- 井原大将・内田ゆず・高丸圭一・木村泰知・江崎浩, 2017「全地方議会会議録の横断検索に向けたデータ収集とデータ構造の検討」『情報処理学会研究報告(Web) (IPSJ Technical Report (Web))』
- 梶原晶・名取良太, 2020, 「大阪市における対立構造の再検討」『地方議会の審議過程—テキスト分析による定量化の試み』関西大学法学研究所研究叢書 63, 57-74.
- 木村泰知, 2021, 「政治情報に関する AI 技術」木村泰知編著、本田正美ほか著『自治体 DX 推進とオープンデータの活用』183-216, 日本経済評論社.
- 増田正, 2016, 「北関東地方における政策課題と地方議会改革: 主要 7 市議会会議録のテキストマイニング分析」『地域政策研究』18, 2・3 号合併号, 33-49.
- 増田正, 2017, 「計量テキスト分析によるわが国地方議会の審議内容を可視化する方法について」『地域政策研究』19 (3), 161-175.
- 増田正, 2018「我が国地方議会における政治・行政関係の計量テキスト分析」『地域政策研究』20(3) 1-19
- 増田正・爲我井慎之介, 2019「地方議会の「会派」は政策集団であるか: 計量テキスト分析による予備的研究」『地域政策研究』22(2): 85-100.
- 松本直樹, 2008「地方議員の図書館への関心に関する予備的考察: 埼玉県市議会の議会会議録分析を元に」『日本図書館情報学会誌』54(1), 39-56.
- 宮坂久美子, 2019「常任委員会における報告事項に関する一考察: 神奈川県議会総務政策常任委員会を事例として」『公共政策志林』7, 141-155.
- 名取良太・岡本哲和・石橋章市朗・坂本治也・山田凱, 2016, 「地方議会データベースの開発と利用」『情報研究: 関西大学総合情報学部紀要』44, 31-42.
- 名取良太・岡本哲和・石橋章市朗, 2020「首長と議会の対立構造と審議過程—ネガポジ分析を用いた検証」『地方議会の審議過程—テキスト分析による定量化の試み』関西大学法学研究所研究叢書 63, 39-56.
- 大山英久, 2007, 「地方議会の公開と会議録をめぐって」『レファレンス』2007(6): 31-46.
- Steyvers and et al., 2011, “Combining Background Knowledge and Learned Topics” (<https://doi.org/10.1111/j.1756-8765.2010.01097.x>)
- 鈴木潔, 2018「自治体議会の司法政治学: 日田市サテライト訴訟と国立市景観訴訟を事例に」年報政治学 政治学 69 (1): 1\_47-1\_69.
- 高丸圭一, 2019, 「地方議会会議録コーパスと地方議会会議録を用いた学術研究の現状」『知能と情報』31(2): 25-33.
- 田中智和, 2020「地方議会における発言は性別によって異なるのか— TF-IDF 分析を用いた検証」『地方議会の審議過程(研究叢書)』関西大学法学研究所, 21-38.
- 土山希美枝, 2018, 「政策議会の政策過程と『争点資源』開発」廣瀬克哉編著『自治体議会改革の固有性と普遍性』, 103-127.
- 内田ゆず・高丸圭一・乙武北斗・木村泰知, 2019「対数尤度比と政治語彙度を用いた議員の議会活動の可視化」『知能と情報』31(2): 662-671.
- 渡部春佳, 2020, 「地方議会議事録分析による話題抽出についての一試論——地方議会・委員会での公の施設「劇場・音楽堂等」に関する議論を事例に」『社会情報学』9 (1): 1-15.
- 渡部春佳・浅谷公威, 2022「地方議会議事録の統合と利活用に関する社会情報学的研究」『電気通信普及財団報告書』
- 横田慎一, 2022「議会の監視機能に着目した公会計情報活用の考察: —与野党会派別の議事に係るテキストマイニング」『会計検査研究』66, 29-46.

(注書き)

本研究の設計や分析は、研究代表者・共同研究者が共同で行ったものであるが、本報告書の文責は研究代表者にある。

(謝辞)

議会議事録の収集にあたっては、早稲田大学大学院基幹理工学研究科・河岡諒さんに多大なご協力をいただきました。本報告は、公益財団法人電気通信普及財団 2021 年度研究助成および、早稲田大学特定課題研究助成費(課題番号 2021C-603)による研究成果です。ここに記して謝意を示します。

### 〈発表資料〉

題名	掲載誌・学会名等	発表年月
横断的な地方議会議事録の収集による政策研究への活用可能性の検討	社会情報学会中国・四国支部 2022 年度第一回研究発表会	2022 年 12 月
地方議会議事録の横断的収集に基づく 公立劇場に関する議論の分析	社会・経済システム学会	2022 年 10 月