

テキスト分析による審議会等の議事の可視化に向けて —在宅医療に関する中央社会保険医療協議会の 議事を例に—

VISUALIZING DISCUSSIONS IN ADVISORY COUNCILS BY TEXT ANALYSIS: A
CASE OF CENTRAL SOCIAL INSURANCE MEDICAL COUNCIL

松岡 広¹・森川 想²

¹ 公共政策修士（専門職） (E-mail:hirom.matsuoka@gmail.com)

² 博士（法学） 東京大学大学院助教 社会基盤学専攻 (E-mail:morikawa@civil.t.u-tokyo.ac.jp)

本稿では、先行研究でこれまで審議会等の議事録について用いられてきた計量テキスト分析の手法を中央社会保険医療協議会（中医協）の議事録に適用し、その有効性を検証するとともに、実用化に向けた課題を適示する。分析の結果、先行研究で用いられている手法は、中医協の議事録の分析に対しても一定の有効性があることが分かった。具体的には、①テーマ分析により、事前に中医協の会長を分析対象データから除去することによって、TFIDF法により、一定の具体性があるテーマを抽出した上で、そのテーマの委員会ごとの変遷、および、各委員の各テーマへの言及の程度を把握できること、②サブテーマ分析により、ある特定のテーマについての委員の主張の特徴を一定程度把握できること、が分かった。

キーワード：テキスト分析、可視化、中央社会保険医療協議会

1. はじめに

公共政策の形成過程の実践においては、政策目標の達成のために、対象集団と要望・期待を相互伝達することが必要になる。そのために設定される重要な場として、審議会等がある。この審議会等は、政策形成における専門知識の導入がその制度目的の一つであるといえるが、一方で、専門知識の導入によって、審議会等の他の重要な目的である、行政の国民参加や公正の確保、利害の調整に失敗するようでは、その存在意義が危ぶまれることになる。こうした状況は、特に利害が対立するアクター同士に、情報の非対称性があるような政策領域において発生しやすいと考えられる。専門知識を導入しつつ、行政の国民参加や公正の確保、利害の調整を実現するために、審議会等における議論を透明化することは、特定の関係者の利益のみが政策に反映されることを防ぎ、より適切な意思決定を行う上で不可欠なことであるといえるだろう。

本研究で扱う中央社会保険医療協議会（以下、「中医協」）は、こうした構造を有し、また、それがゆえに、透明性を高める目的でその組織についてこれまで幾度かの改変が加えられてきた審議会である。一方で、中医協は他の審議会等に比べても高い頻度で開催されており、そこにおいて議論される議事も膨大かつ広範な内容を扱う

ものであり、市民社会がその議事を逐一確認してその妥当性を検討することは極めて難しい。しかし、以下に紹介するように、近年、テキストを量的に分析することにより、議論の構造や内容を要約する試みが行われており、そのためのツールも開発されてきている²⁾。こうした手法を審議会等における議事録に適用することで、行政がこれまでの議論を適切に要約して、より重要な論点を抽出するのに役立てたり、市民社会が、審議会等における議論の内容を簡易に確認したりするのに用いることはできないだろうか。本稿では、他の審議会等について用いられてきた手法を中医協の議事録に適用し、その有効性を検証するとともに、実用化に向けた課題を適示する。

2. 先行研究

医療政策の政策過程に関して、実務上も重要であり、なおかつ分析上も高い関心を持たれてきたのは、診療報酬の改定に関する政治過程であろう。文献⁴⁾は、診療報酬改定の重要な性質の一つとして、「改定の内容について合意を得るには、ともかくも強力な利害関係が生じることを考慮に入れれば、それは政府にとって困難な政策課題となっていくのである」と述べている。診療報酬の改定は本研究で分析対象とする中医協をその場として実施

されており、そこでの重要なアクターとしては、厚生(労働)省、日本医師会、政党、健康保険連合組合などがある⁵⁶⁾。その決定プロセスについては文献 8)9)、その歴史的背景については文献 10)に詳しい。診療報酬の改定以外の政治過程についても、1972 年の医事基本法を題材に、政策過程分析を行った文献 11)は、行政部門としての厚生労働省と、政治的アクターとして利益団体(日本医師会)、政党、世論との関係を分析している。本研究で用いたテキスト分析を用いた医療政策関連の研究としては、新聞記事情報にテキスト分析を適用し、医療分野の課題を抽出したのものとして文献 12)があるが、中医協の議事録を素材として分析を行い、議論の内容を分析しようとする試みは管見の限り存在しない。

一方、本研究が分析の素材として扱う審議会等における議事録であるが、その多くは情報公開の一環として公開されており、政策過程の研究において重要な資料となっている。公的な議事録に対して定量的な分析を施した先行研究としては、討議参加者の発言に「賛否」「論証性」「厳密性/適切性」「肯定的態度/否定的態度」の4つのファセット情報を付加した上で参加者間の認識の不一致や意見対立を抽出したもの¹³⁾、発言の類似性により議論を構造化したもの¹⁴⁾、テキストマイニングにより発言者の意見を抽出し発言の推移を抽出したもの¹⁵⁾、キーワード間の関連性やキーワードと発言者の関連性を多重対応分析やクラスター分析を用いて分析したもの¹⁶⁾などがある。

しかし、政策過程の研究という観点からは、審議会においてどのようなテーマが主に話し合われたか、また、各委員がどのテーマに対して発言をしていたか、を把握することが肝要である。この観点において重要な先行研究として、岩見らが実施した、淀川水系流域委員会本委員会の議事録を対象として、話し合われたメインテーマとサブテーマ、それらの変遷、サブテーマを介した委員間の関係性、委員間の意見の協調・対立関係を把握したもの¹⁷⁾¹⁸⁾¹⁹⁾²⁰⁾がある。本研究では、先行研究において岩見らが開発した手法を、国の審議会の議事録に適用し、その有効性を検証するとともに、実用化に向けた課題を適示することを目的とする。

3. 分析手法

文献 18)においては、テキストマイニングを用いて議論内容を把握し提示するために、分析対象とする語の選定手法を含む分析手法を開発し、その手法によって、委員会において話し合われたテーマとその変遷の把握(以下、「テーマ分析」)がなされている。また、文献 19)では、テーマ分析により抽出された特定のテーマについてサブテーマを抽出し、委員ごとの各サブテーマに対する

言及の程度を分析(以下、「サブテーマ分析」)することにより、各委員の主張やその類似点・相違点など、委員間の関係性の把握を試みている。

本研究では、これらの先行研究をもとに、厚生労働省の審議会である中央社会保険医療協議会(中医協)の議事録を対象として以下の分析を行う。

- (1) テーマ分析: 語の出現の偏りに着目した選定手法により分析対象とする語を選定、それをもとに話し合われたメインテーマを特定し、その変遷を把握する。
- (2) サブテーマ分析: 特定のメインテーマについて、発言した委員による語の偏りに着目した選定手法により分析対象語を選定、それをもとに話し合われたサブテーマを特定し、各委員による言及の程度を把握する。

なお、分析用ソフトウェアとしては、自然言語処理と多変量解析の両方を実行可能で、研究に広く利用されているフリー・ソフトウェアである樋口耕一氏の KH Coder を用いる³⁴⁾。

3.1. 分析対象

中医協は、健康保険者等を代表する7名の委員(「支払側委員」)、医師・歯科医師・薬剤師を代表する7名の委員(「診療側委員」)、公益を代表する6名の委員(「公益委員」)の計20名の委員から構成されており、診療報酬の体系、個別の点数について審議決定を行っている。中医協で審議されるアジェンダは、1) 新規の医薬品の保険収載とその価格、新規の医療機器、医療材料、臨床検査の価格の決定、2) 診療報酬改定の結果についての調査に関する案件、3) 診療報酬体系における制度に関する案件、4) その他、の4つの類型が挙げられる⁹⁾が、本研究では、支払い側委員と診療側委員の意見の対立が明確に見られると考えられる、3) 診療報酬体系における制度に関する案件に着目する。そのなかでも、2014年度の診療報酬改定に向けては在宅医療について特に意見が対立し白熱した議論が展開されたことが指摘されている⁹⁾ことから、在宅医療について審議された部分を分析対象とする。具体的には、以下の在宅医療に関する議事録である。

・在宅医療に関する議事録: 中医協総会第237回(2013年2月13日)、第242回(同年5月29日)、第244回(同年6月26日)、第252回(同年10月23日)、第253回(同年10月30日)の議事録のうち、在宅医療に関する部分

テキスト	発言者	日付
医療課長でございます。	宇都宮医療課長	20130213
資料「中医協総-5」をごらんいただけますでしょうか。「在宅医療」	宇都宮医療課長	20130213
3枚目は、社会保障・税一体改革の図、大綱の抜粋が出てござい	宇都宮医療課長	20130213
7枚目の「平成24年度診療報酬改定の基本方針のポイント」は、	宇都宮医療課長	20130213
8枚目は、基本方針です。	宇都宮医療課長	20130213
9枚目は、答申書です。	宇都宮医療課長	20130213
10枚目は、答申書附帯意見丸1です。	宇都宮医療課長	20130213
11~13枚目に、附帯意見の中で在宅医療についての指摘がござ	宇都宮医療課長	20130213
いほど御説明しますが、基本的には、この附帯意見に基づきま	宇都宮医療課長	20130213
14枚目は「2. 在宅医療を取り巻く現状について」でございます。	宇都宮医療課長	20130213

Fig. 1 議事録データの例

3.2. データの前処理

(1) 議事録データの整備

KH Coder には、分析対象のテキストとその外部変数（例えば、各テキストの発言者）とのリスト（Excel 形式）をテキスト形式に自動的に変換して、外部変数とともに読み込む機能がある。そのため、今回の分析にあたっては、インターネット上で公開されている中医協の議事録を、Fig. 1 のように、テキストにその発言者と日付とが紐づいたリスト形式に整備した。

この際、リストの各行のテキストは1つの「段落」として KH Coder に読み込まれ分析の単位となるが、これは、オリジナルの議事録の各段落に対応させた。なお、議事録のうち、中医協会長の発言は、各委員の発言を促すなど、議論の実質に関係ないものがほとんどであったことから、今回の分析対象から除外した。

(2) 強制抽出する語と使用しない語の指定

KH Coder では日本語データ中から語を取り出すために使用する外部プログラムとして茶筌を使用しているが、文章から自動的に語を取り出す際に、「在宅医療」という言葉が「在宅」と「医療」という2つの語として認識され、意図通りの抽出が行われない場合がある。こうした専門的な複合語を抽出するためのアルゴリズムのうち、KH Coder の「複合語の検出」機能では、TermExact と呼ばれる専門用語抽出モジュールを利用して複合語を抽出している²¹⁾。本研究では、この機能を用いて、複合語として抽出した方が良いと考えられる語の組み合わせを洗い出し、スコアの高い複合語 246 語を「強制抽出する語」として指定した。例えば、4000 以上のスコアを持つ複合語として、在宅医療、訪問看護、医療機関、医療保険、訪問診療、訪問看護ステーション、介護保険、診療報酬の 8 つが抽出された。

一方、「スライド」や「ページ」など、議論の実質に関係ない語も抽出されてしまうことから、出現頻度の高い語のうち議論の実質に関係ないと考えられる語 34 語を「使用しない語」として指定した。

(3) 前処理の実行

上述の指定を行った後、KH Coder で「前処理の実行」

(分析対象ファイル内の文章から語を切り出し、その結果をデータベースとして整理)を行った。この際、岩見らの先行研究¹⁷⁾¹⁸⁾¹⁹⁾²⁰⁾に倣い、分析に使用する語を、名詞の一般（漢字を含む2文字以上の語）、サ変接続の名詞、固有名詞、組織名、地名、複合語（KH Coder 上では「タグ」と表示）のみとした。

前処理の結果、在宅医療の議事録については、1,344 の段落に対して 4,038 の文が認識され、分析に使用する語としては、1,692 種類、延べ 13,254 語が抽出された。

3.3. テーマ分析

岩見らの先行研究¹⁷⁾¹⁸⁾¹⁹⁾²⁰⁾では、テーマ分析における分析対象語選定の手法として、以下の3つが挙げられている。基本的な発想としては、(1)議事録に多く出現する語を分析対象語として選定する、(2)特定の場所に偏って出現する語を分析対象語とする、の二つの基準のバランスによって、出現回数及び出現段落の偏りから構成されるスコアを算出し、その大きい順に分析対象語を選定することになる。

- (a) TFIDF 法：語 w の出現回数 (TF: Term Frequency) と、出現段落の偏り (IDF: Inverse Document Frequency) の 2 つの基準によって指標を算出

$$TFIDF(w) = TF(w) \times \log\left(\frac{D}{DF(w)}\right)$$

(TF(w)は語 w の出現回数、D は全段落数、DF(w)は語 w の出現段落数)

- (b) 変動係数法：各会合における、語 w の出現段落数の変動係数（標準偏差と平均の比）を算出

$$CV(w) = \frac{\sigma_{DF}(w)}{DF(w)}$$

(分子は語 w の各会合の議事録における出現段落数の標準偏差、分母は同出現段落数の平均値)

- (c) DFIMF 法：語 w の出現段落数 (DF: Document Frequency) と、出現委員会の偏り (IMF: Inverse Minutes Frequency) によって指標を算出

$$DFIMF(w) = DF(w) \times \log\left(\frac{M}{MF(w)}\right)$$

(M は委員会の総開催回数、MF(w)は語 w が出現した委員会の回数)

id	患者	病院	薬局	データ	ステーション	状況	資料	実績	要件	論点
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	1	0	0
14	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0
18	0	1	0	0	0	0	0	0	0	0
19	0	1	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	1	0	0	0

Fig. 2 クラスタ分析に用いられる集計表（『文書×抽出語』表の一部）

本研究では、前処理で抽出した1,692種類の語のうち、5回以上出現していた540語について、各手法を用いてスコアを算出し、上位100語を分析対象語として選定した。なお、文献18)では、TFIDF法では、議論内容に関係のない、会議運営のために繰り返し発言された語が選定語に多く見られ、テーマの特定が困難となるため、テーマ分析の手法としてはDFIMF法に劣る、とされている。

次に、それぞれの方法によって抽出された分析対象100語を、KH Coderの階層的クラスタ分析の機能を用いて、10のクラスタに分類した。KH Coderの階層的クラスタ分析では、Fig. 2に示されるような、それぞれの語の各段落での出現数が入力された集計表（KH Coderでの呼称は『文書×抽出語』表）が用いられ、出現パターンの似通った語が同じクラスタに分類される（図中のidが段落番号に対応する）。クラスタ間距離の定義としてはJaccard係数、ユークリッド距離またはコサイン距離の3種類、クラスタ化法としてはWard法、群平均法、最遠隣法の3種類から選択できるが、本分析ではJaccard係数とWard法を選択した。

そして、各クラスタに分類された語群から、それらに共通するテーマを見出して名付け、分類された語群が特定のテーマに関するキーワードであると考えた。その上で、あるクラスタに含まれる語が延べ2回以上出現している段落を、そのクラスタに対応するテーマについて話された段落とみなして、各委員会、および各委員の発言における各テーマの出現頻度を把握した。（したがって、一つの段落に複数のテーマが対応することや、対応するテーマが存在しない段落も存在する。）

3.4. サブテーマ分析

サブテーマ分析においては、テーマ分析において抽出されたメインテーマから1つを選択し、議事録のうちそのテーマについて話されたとみなされる部分（すなわち、そのテーマに対応するクラスタ語群に含まれる語が2

回以上出現する段落の集合）に対し、テーマ分析と同様の分析を行った。その際、分析対象語の選定においては、岩見らの先行研究¹⁹⁾に倣い、発言委員の偏りに着目した選定手法であるDFIPF（Document Frequency Inverse Personal Frequency）を用いた。

$$DFIPF(w) = DF(w) \times \log\left(\frac{P}{PF(w)}\right)$$

（Pは分析対象部分において発言した総委員数、PF(w)は語wを発した委員数）

また、分析対象の段落数、語数が少なかったことから、分析対象語を約50語、クラスタ分類数を7とした上で、各委員の発言内の各サブテーマの出現頻度を把握した。

4. 分析結果

4.1. テーマ分析

TFIDF法、変動係数法、DFIMF法それぞれによって、分析対象語100語を抽出し、階層的クラスタ分析を適用した結果は、Fig. 3~5の通りである。第1列に示したデンドログラムでは、クラスタ間の距離に応じてどのクラスタ間に近接性がみられるかを示している。各クラスタに分類された語が第3列に挙げられているが、このクラスタに分類された語群をもとに、それぞれのクラスタに対して、テーマ名を与えたのが第2列である。TFIDF法で分析対象語を抽出し分析した結果（Fig. 3）では、「在宅療養支援診療所/病院（在支診/病）の評価」「訪問看護」「保険給付」「地域に応じた在宅医療」「在宅医療と病院との連携」「薬局における在宅業務」「自宅での看取り」、変動係数法で分析対象語を抽出し分析した結果（Fig. 4）では、「高齢化、看取りの状況」「無菌調剤」「在宅医療における薬剤」「看護師のキャリアパス」「患者紹介ビジネスの規制」「中核的な訪問看護ステーションの評価」「在宅歯科医療」「サービス付き高齢者向け住宅」、DFIMF法で分析対象語を抽出し分析した結果（Fig. 5）では、「無菌調剤」「患者紹介ビジネスの規制」「看護師のキャリアパス」「機能強化型在支診/病」「サービス付き高齢者向け住宅」「在宅療養あんしん病院登録システム」「訪問看護ステーション」「看取りの状況」「薬局における在宅診療」といったテーマ名を与えた。（クラスタに分類された語からテーマを特定することができなかったクラスタについては、「特定不能」としている。）

さらに、各手法により抽出されたテーマ（テーマが特定不能であったクラスタを除く）の各会合、および、各委員の発言における出現頻度を色の濃さで示したのが、Fig. 6~8である。なお、各図において、枠内の数字は各会合において、そのテーマが出現した段落の割合（百分

デンドログラム	テーマ名	語
[Dendrogram]	在支診/病の評価	実績 機能強化 在支病 最後 課題 評価 論点 連携 医療機関 要件 強化
	訪問看護	規模 訪問看護 ステーション 大規模化 事業所
	保険給付	給付 介護保険 医療保険
	地域に応じた在宅医療	都道府県 医師会 日本 役割 地域包括ケア 在宅医療 推進 中心 仕組み 地域 診療所
	在宅医療と病院との連携	京都 奈良 病院 入院 一 確保
	特定不能	実施 歯科訪問診療 増加 グラフ 事例 診療 歯科 訪問診療 算定 往診 回数 医師 施設 訪問 かかりつけ
	薬局における在宅業務	薬剤師 薬局 在宅業務
	自宅での看取り	自宅 家族 高齢者 介護 医療
	特定不能	議論 中核 現場 改定 診療報酬 検証部会 報告 医療 理解 説明 調査 資料 検討 ケース 具体 データ 指標
	特定不能	ニーズ サテライト 管理 体制 加算 対応 状況 事務局 質問 方向 方向性 意見 機能 保険医療機関 制度 紹介 在宅 患者 意味 提供 情報

Fig. 3 階層的クラスター分析の結果 (TFIDF 法)

デンドログラム	テーマ名	語
[Dendrogram]	高齢化、看取りの状況	世帯 高齢化 我が国 フランス 外国 老健 入所 介護施設 ケア付き 住宅 死亡 場所
	無菌調剤	無菌調剤 共同利用 無菌
	在宅医療における薬剤	電解 製剤 点滴 注射
	看護師のキャリアパス	キャリア パス
	患者紹介ビジネスの規制	療担規則 患者紹介 ビジネス 行為 利益
	特定不能	あんしん 在宅療養 特定保険医療材料 費用 材料 衛生材料 管理指導 地図
	中核的な訪問看護ステーションの評価	拠点 中核 資格
	在宅歯科医療	口腔 歯科 歯科医療 在宅療養 在宅療養 歯科診療所 写真 在宅 歯科医療 口腔ケア 介護者 研究 集合 扱い 呼び出し 看護
	サービス付き高齢者向け住宅	運営 医療法人 サービス付き高齢者向け住宅 搬送 サービス付き高齢者向け住宅 法人
	特定不能	在宅療養支援病院 担当患者 方策 ネットワーク 管理業務 共有 チーム 医療系 スタート 記憶 不適切事例 排除 減少 指示書 要因 看護職員 常勤 24年度改定 軽減 見方 届け出 居宅介護支援事業所 介護支援専門員 医学総合管理科 計画 予測 年間 フリー 保険者 基本診療料 固定 人材 訪問歯科診療 規則 支援診療所 表記 医療区分 記録 メン 講師 サテライト 診療料 指標 規模

Fig. 4 階層的クラスター分析の結果 (変動係数法)

デンドログラム	テーマ名	語
[Dendrogram]	無菌調剤	製剤 電解 無菌調剤 無菌 共同利用
	患者紹介ビジネスの規制	ビジネス 患者紹介 利益 禁止 紹介料 紹介 療担規則 保険医療機関 誘導
	看護師のキャリアパス	キャリア パス
	機能強化型在支診/病	サービス付き高齢者向け住宅 グループ 法人 趣旨 点数 緊急往診 実績 機能強化 在支病
	サービス付き高齢者向け住宅	サービス付き高齢者向け住宅 サ高住 医療法人
	在宅療養あんしん病院登録システム	登録 あんしん 在宅療養
	訪問看護ステーション	看護職員 常勤 職員 事業所 ステーション 規模 大規模化 分析 サテライト ネットワーク
	看取りの状況	減少 増加 推計 老健 ケア付き 我が国 老人ホーム 死亡 自宅
	特定不能	介護支援専門員 居宅介護支援事業所 中核 届け出 業務 運営 搬送 在宅療養支援診療所 在宅療養支援病院 入院 中小病院 往診 一定 在宅専門 かかりつけ 医 記録 情報提供 共有 口腔 介護者 在宅歯科医療 写真 歯科 医療 在宅療養 歯科診療所 歯科訪問診療 訪問 歯科 指導 支援診療所 意識 給付 不適切事例 排除 事例 業者 印象 看護 呼び出し 軽減 検証 速報 夜間 メン 医政局 薬剤管理 医療区分 医学総合管理科
	薬局における在宅業務	点滴 衛生材料 特定保険医療材料 費用 薬局 在宅業務 薬剤師 チーム 医療

Fig. 5 階層的クラスター分析の結果 (DFIMF 法)

率) を表している。なお、この頻度によって、Fig. 2 と同様の集計表を作成することができ、これに基づいて階層的クラスター分析を行った結果も併せて示している。これにより、どのテーマが各会合、および、各委員において類似した出現傾向にあるかを見ることができる。また、委員名の前についている数字は、1 が支払側委員、2 が診療側委員、3 が公益委員、4 が専門委員、5 が事務局を表している。

TFIDF 法、変動係数法、DFIMF 法の3つの方法を比較する上で、それぞれの方法で特定されたテーマが、各会合や各委員の発言における出現頻度の面で十分なものかどうか、また、名付けたテーマが実際の議論の内容と対応しているかどうか重要な基準となると考えられる。

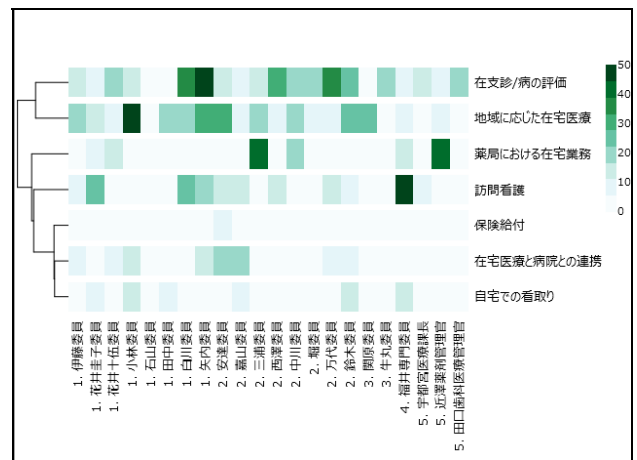


Fig. 6 テーマの各会合における出現頻度 (上) と各委員の発言における出現頻度 (下) (TFIDF 法)

この観点から TFIDF 法、変動係数法、DFIMF 法の3つを比較すると、TFIDF 法では、より抽象度が高く出現頻度も高い語 (例えば、「最後」、「課題」、「論点」など) が分析対象語として選定される傾向が大きい一方、変動係数法では、より具体的ではあるが出現頻度が低い語 (例えば、「無菌調剤」、「療担規則」など) が選定される傾向が大きく、DFIMF 法は両手法の間であることが分かる。その結果、TFIDF 法においては、抽出されたテーマの出現頻度は概ね高い。他方、変動係数法では各テーマの出現頻度が極めて低く、DFIMF 法でもほとんどのテーマの各会合での出現頻度が10%未満と低かった。

実際の議論との対応を見ても、例えば第3回会合で議論された「在支診/病の評価」や、第1回会合で議論された「地域に応じた在宅医療」について、TFIDF 法においては、実際に活発に議論がかわされたテーマを抽出できている。これに対し、変動係数法では、クラスター分析の結果の段階においてすでに、こうしたテーマ名をクラスターに与えることが困難となっている。DFIMF 法では、

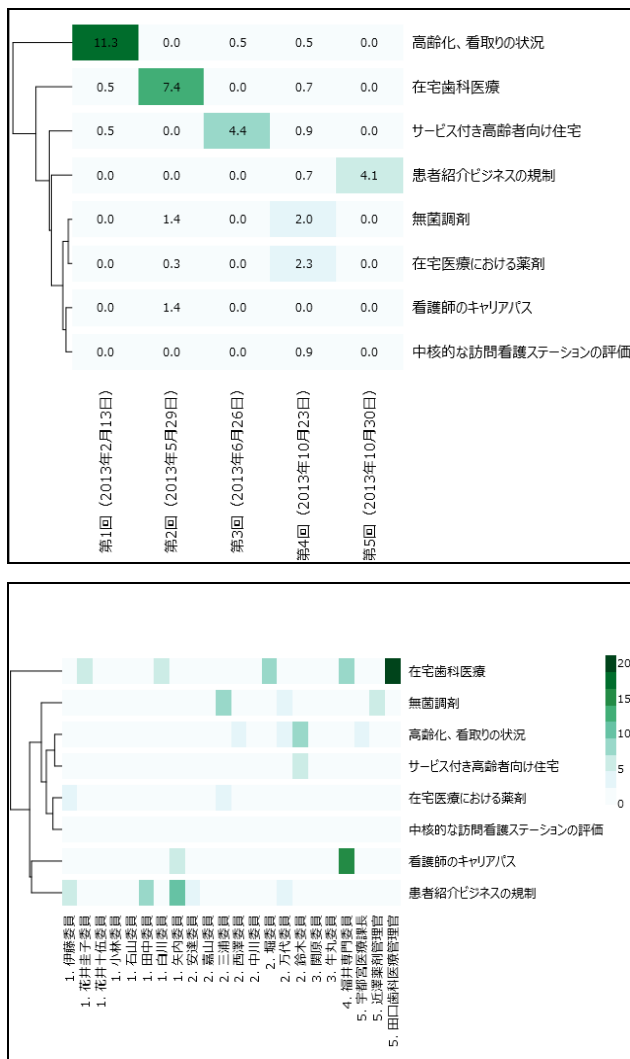


Fig. 7 テーマの各会合における出現頻度 (上) と各委員の発言における出現頻度 (下) (変動係数法)

実際の議論との対応においても TFIDF 法と変動係数法の間期的な結果が得られているが、例えば第1回会合の議事録におけるテーマの出現割合は、「看取りの状況」に多くなってしまっており、「地域に応じた在宅医療」について捉えることができていない。このように、変動係数法や DFIMF 法では、TFIDF 法のように、活発に議論がかわされたテーマを抽出することができなかった。

先述のように、文献 18)では、TFIDF 法では、議論内容に関係のない、会議運営のために繰り返し発言された語が選定語に多く見られ、テーマの特定が困難となるため、テーマ分析の手法としては DFIMF 法に劣る、とされている。本分析でそれと異なる結果が得られた理由としては、①データの预处理の段階で中医協会長の発言を除外したため、会議運営のために繰り返し発言された語がそもそも分析対象から外れていたこと、②岩見らの先行研究¹⁷⁾¹⁸⁾¹⁹⁾²⁰⁾(80 回以上の委員会の議事を対象)に比べ、委員会の開催回数をはるかに小さく、重要な単語で

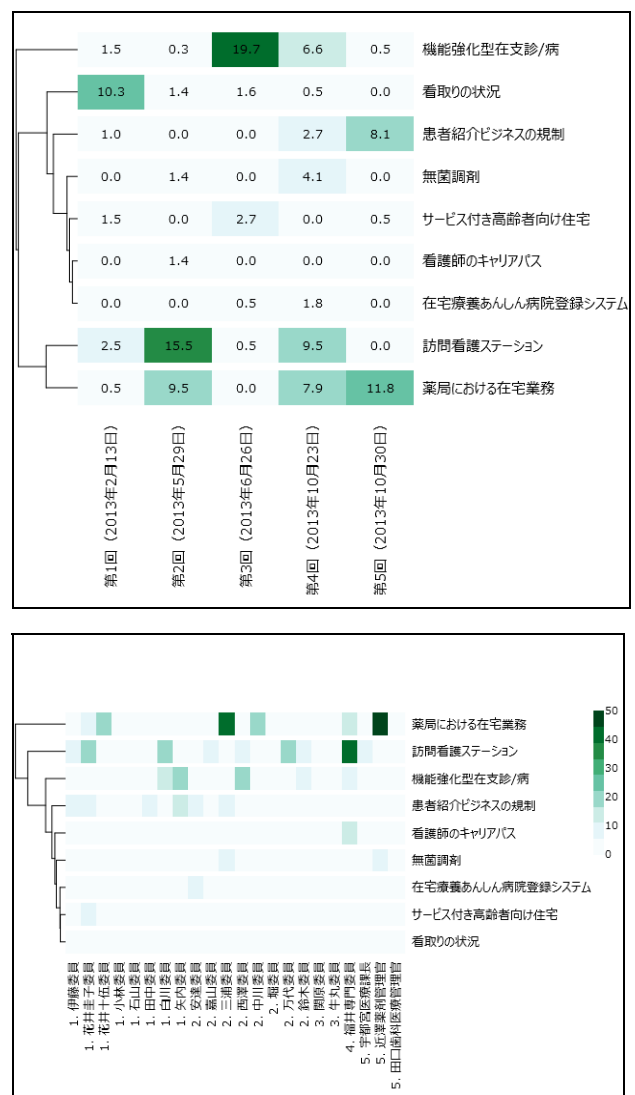


Fig. 8 テーマの各会合における出現頻度 (上) と各委員の発言における出現頻度 (下) (DFIMF 法)

も 5 回全ての会合の議事録に出現していたために、DFIMF 法によって分析対象として選定されなかったこと、の二点が考えられる。

いずれの違いにせよ、今回の分析においては、TFIDF 法の方が、他の手法よりも有効性が高いという結果が得られた。

4.2. サブテーマ分析

続いて、TFIDF 法で抽出されたテーマのうち、「地域に応じた在宅医療」のテーマについてサブテーマ分析を実施した。先に紹介したように、岩見らの研究¹⁹⁾では、特定のテーマについてサブテーマ分析を行うことによって、アクター間の利害の対立構造を試みている。今回、「地域に応じた在宅医療」のテーマとして選択した理由は、文献 9)において、在宅医療に関して支払側委員と診療側委員との間で白熱した議論がかわされたテーマとして、「在宅医療の提供のあり方」(都市部で増加している

在宅専門医療機関を容認するのか、各地域で「かかりつけ医」機能を担ってきた一般の診療所が在宅医療の担い手として評価するのか)が挙げられており、委員間の主張の対立構造を分析する上で適したテーマであると考えたためである。例えば、「地域に応じた在宅医療」のテーマに分類された発言としては、以下がある。

“さらに地域においては、医療機関である、診療所、有床診療所、中小病院、そして郡市区医師会、そういったところこそ多機能化して、24時間対応する巡回型の訪問看護ステーションと連携して、地域包括ケアを支えていくべきだと考えております。そういう意味では、有床診や中小病院の入院機能もその1つだと考えられます。そういった形で、日本型の既存資源を活用した、超高齢社会を支えていく仕組みをつくっていくのがよろしいのではないかと考えております。そういう意味では、少し方向性が違うのではないかという気が、全体としていたしました。” (2013年10月23日、診療側委員：鈴木委員の発言)

“今の安達先生のお話でございますが、在宅医療を担う医師は、かかりつけ医的な方が理想であろうというのは、十分に理解しますし、私もそういうふうになります。ただ、現実問題、少数の医師で在宅医療を担い、かつ24時間、緊急対応もというのは、医師の負担が大き過ぎるということで、いろいろな工夫をしていこうというのが、今回あるいは前回の考え方だと思っております。したがって、在宅専門の診療所がいかに、悪いかというのは、もう少し突き詰めて議論をしなければいけないとは思いますが、今、在宅医療は、歯科、薬局を含めて、体制をつくっていこうという段階でございますので、余り固定的な考え方に捉われないほうがいい。トライアルみたいなこともあるでしょうし、いろいろな形態が地域によってあり得ると考えております。” (2013年10月30日、支払側委員：白川委員の発言)

議事録のうち、「地域に応じた在宅医療」について議論された部分について、DFIPF法によって、分析対象語53語を抽出(49番目から53番目の語のスコアが同じであったため)し、階層的クラスター分析を適用した結果は、Fig. 9の通りである。Fig. 3~5の場合と同様、各クラスターに分類された語をもとに、それぞれのクラスターに対して、「医師会の在宅医療の取り組み」「中小病院と有床診療所の入院機能」「訪問看護」「実績の評価」「在宅医療専門医療機関の外来応需」「在宅医療専門医療機関とかかりつけ医機能」「薬剤師の活用」のテーマ名を与えた。

デンドログラム	テーマ名	語
	医師会の在宅医療の取り組み	京都 機構 システム 紹介 確立 取り組み 医師会 日本 都道府県 支援
	中小病院と有床診療所の入院機能	入院 有床診療所 中小病院
	訪問看護	訪問看護 ステーション 地域包括ケア 役割 連携 地域 医療機関 中心 在宅医療 推進 方向性 仕組み 医療
	実績の評価	論点 評価 強化 改定 実績
	在宅医療専門医療機関の外来応需	資料 課題 要件 応需
	在宅医療専門医療機関とかかりつけ医機能	議論 かかりつけ医 在宅専門 患者 診療所 病院 在宅 理解
	薬剤師の活用	地域医療 活用 検討 薬局 薬剤師 指摘 ケース 基本 医師 対応

Fig. 9 階層的クラスター分析の結果 (DFIPF 法)

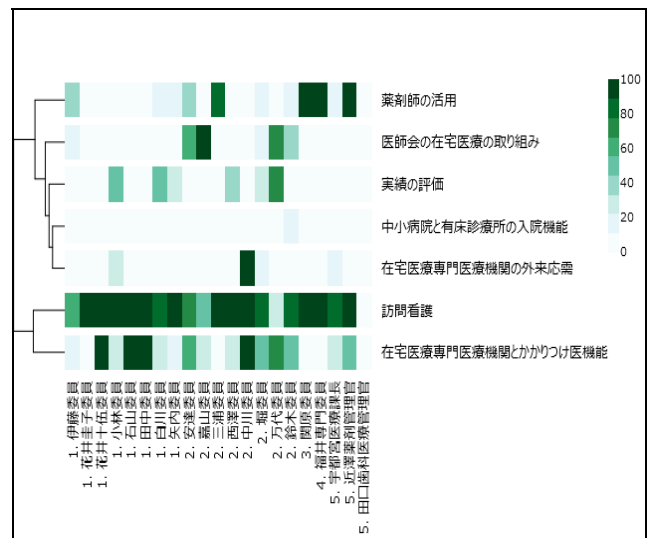


Fig. 10 サブテーマの各委員の発言における出現頻度

また、抽出された各サブテーマの各委員の発言における出現頻度を色の濃さで示したものがFig. 10である。Fig. 6~8と同様、この頻度によって、Fig. 2と同様の集計表を作成することができ、これに基づいて階層的クラスター分析を行った結果も併せて示している。また、委員名の前に付いている数字は、1が支払側委員、2が診療側委員、3が公益委員、4が専門委員、5が事務局を表す。

Fig. 10からは、委員間の対立構造として、診療側委員(2で示される)が「医師会の在宅医療の取り組み」について他の委員よりも多く言及していることが分かる。これは、先の引用のように、診療側委員が、「かかりつけ医」機能を果たしている一般診療所が中心となって今後の在宅医療を推進していくべきとの主張をしていたことによるものと考えられる。他方、支払側委員(1で示される)がどのような論拠を主に掲げていたかは本分析からは不明確であり、それを把握するためには、彼らが比較的多く言及していたサブテーマ(「訪問看護」や「在宅専門医療機関とかかりつけ医機能」など)について個々の発言を見ていく必要がある。

支払側委員の発言を見ると、以下のように、診療側委員の主張が、一般診療所での在宅診療により診療報酬を

配分すべき、という議論に性急に結びついてしまうことの懸念と、在宅医療を担うアクターをより幅広く捉えて、慎重な議論を求める声があることがわかる。

“診療報酬の2割ぐらい紹介料を払う例が出ておりまして、実際には何割払っているか、我々は承知しておりませんが、2割払ってももとがとれるというのは、診療報酬が高過ぎるということになりかねない。もう一つは、2割分を払ったのを取り返すために、過剰な診療をしているのではないかと疑われるということでございます。”

(2013年10月30日、支払側委員：白川委員の発言)

“まだ主治医機能について一生懸命構築している時期に、外来を求めるかどうかという議論は、まだ必要なのではないか。それ以上に主治医機能の充実なり、方向なりが、本当に固まった後に出てくる新しいテーマだと思っております。したがって、現段階では、在宅医療、訪問医療を含めて、主治医、かかりつけ医機能をどうやって充実させていくかというふうに議論を集中させたほうが、得策ではないかというのが、私の意見です。”

(2013年10月30日、支払側委員：石山委員の発言)

“今回提案された内容で、地域包括ケアにおいて中核的な役割を果たすということ、それから、先ほど課長のお話の中に、医療と介護の橋渡し役を果たす、あるいは居宅介護支援事業所としての役割もというところ、今回提案されている機能強化型の訪問看護ステーションの担う範囲というの、一体どういうふうになるのかというのが、もう少し提示されて、イメージが湧くようなものがほしいと思います。”

(2013年10月23日、診療側委員：花木圭子委員の発言)

このように、診療側委員が在宅医療専門医療機関について批判的であったのに対し、支払側委員は在宅医療専門医療機関を認めることも在宅医療推進のための一つの可能性であることを留保しており、この点では、両者の意見は対立していたといえる。一方で、診療側委員は明確に従来の診療所が担ってきたかかりつけ医機能をより評価すべきと主張していたのに対し、支払側委員からは在宅医療を担うべきアクターについて明確な提案がなされなかった。この意味では、第253回(2013年10月30日)の会合を森田会長が「1号側、2号側で活発な議論がされたと思います。ただ、基本的な方向はそれほど違わないようにも思いました」とまとめているように、その対立構造は先鋭化したものではなかった。このこと

が、サブテーマ分析において支払側委員の論拠の抽出が難しかった要因と考えられる。

5. まとめと課題

本研究から、岩見ほかの先行研究¹⁷⁾¹⁸⁾¹⁹⁾²⁰⁾で示された手法は、中医協の議事録の分析に対しても一定の有効性があることが分かった。具体的には、①テーマ分析については、事前に中医協の会長を分析対象データから除去することによって、TFIDF法により、一定の具体性があるテーマを抽出した上で、そのテーマの委員会ごとの変遷、および、各委員の各テーマへの言及の程度を把握できること、②サブテーマ分析により、ある特定のテーマについての委員の主張の特徴を一定程度把握できること、が分かった。

他方、以下の3点は本研究では十分にできておらず、今後検討すべき課題として挙げられる。

①分析対象議事録の性質と分析手法との関係

岩見らの先行研究¹⁸⁾では、テーマ分析における分析対象語の選定手法としてDFIMF法の優位性が指摘されていたが、本分析ではTFIDF法の方がDFIMF法よりも有意義なテーマを抽出することができた。この要因として、先に、分析対象から中医協会長の発言を除外したこと、議事録に含まれる委員会の開催回数が少なかったことの2点を挙げたが、今後本分析の実用化を考える上では、どのような議事録に対してどのような分析手法が適しているかをさらに検討していく必要がある。

また、ここでの分析対象である公開された議事録は、各委員の修正が入ったものであるほか、議事の進行自体を通じ事務局や会長の立場が反映される可能性がある⁹⁾。前者については、公開のために修正された文言こそ、それぞれの立場を示すものであるという解釈も可能であるが、削除された発言については補足できないという問題がある。また、後者については、今回のような分析結果全体を、事務局や会長交代前後比較することで、捕捉することが可能かもしれない。

②分析対象語数、クラスター数設定の限界

本研究では、テーマ分析においては分析対象語数を100、クラスター数を10とし、サブテーマ分析においては分析対象語数を約50、クラスター数を7とした。どのような設定が、議論の内容を把握するのに適しているかを一概に議論することは難しい。例えば、分析対象語数の設定については、語数を先に設定するのではなく、指標の閾値を先に設定し、その閾値を超えた語を選定する、という手法も考えられる。今回のテーマ分析において、

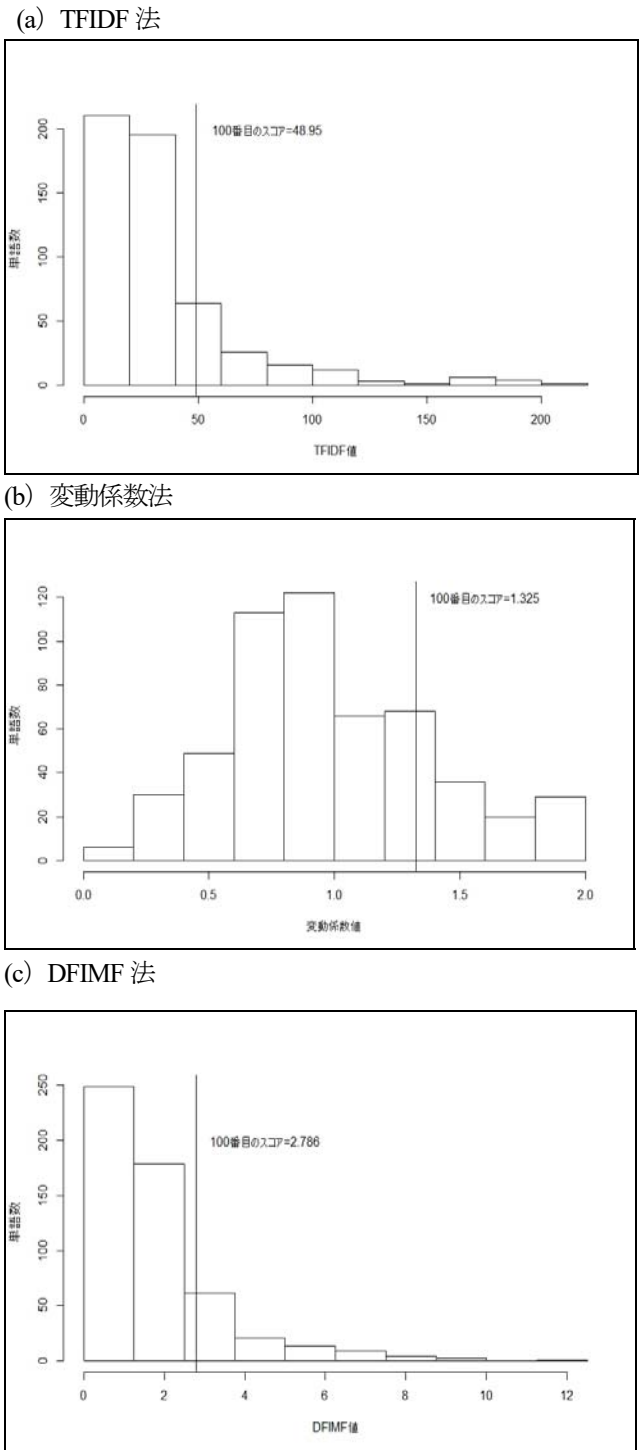


Fig. 11 テーマ分析で用いた各手法でのスコアの分布

各手法により算出された指標（スコア）上における単語数のヒストグラムと、100番目の単語のスコアを Fig. 11 に示す。この図に即していえば、スコアの高い単語を分析対象とすべきであるから、ヒストグラムの右に分布している単語のどこまでを分析対象とするかの基準として、例えば、上から100目までを分析対象とする方法や、分布の形状からスコアの閾値を定め、それ以上の単語を分析対象とする方法が考えられる。今回に関していえば、

TFIDF 法と DFIMF 法では、出現頻度が急激に下がるところを基準とした場合の閾値のスコアと、100番目の単語のスコアが似た値をとるため、分析対象語数の設定をどちらにした場合も大きな違いはないが、変動係数法では、スコア上での単語数の分布を踏まえた閾値設定という観点からは、100番目までの単語を分析対象とすることに、客観的な理由は見出しにくいことになる。

③委員間の意見の協調・対立関係の把握

サブテーマ分析では、各委員の論拠を一定程度把握できることが分かったが、委員同士がそのテーマについて協調しているか対立しているかまでは把握できなかった。例えば、サブテーマ分析において抽出されたサブテーマ「在宅医療専門医療機関とかかりつけ医機能」については、支払側委員、診療側委員ともに多く言及しているが、診療側委員が在宅医療専門医療機関について批判的で従来の診療所が担ってきたかかりつけ医機能をより評価すべきと主張する一方で、支払側委員は在宅医療専門医療機関を認めることも在宅医療推進のための一つの可能性であることを留保していた点では、意見は対立していた。

この点について、KH Coder の「文書検索」機能を用いて、今回行ったように当該サブテーマに分類された発言の内容を確認することで定性的に意見の協調・対立関係を把握することは可能であるが、定量的に把握し可視化する手法の開発が今後の検討課題である。なお、文献20)では、委員間の発言の応酬に関する情報（ある委員に対して他の委員が応答した回数）をもとにネットワークグラフを描画することによって委員間の意見の協調・対立関係を可視化することに成功しているが、テキストデータ以外の情報に依拠しない手法の開発が実用化に向けては必要であるといえる。

本研究は、以上に挙げたような課題を残すものの、審議会において話し合われたテーマを抽出し、かつ各委員の主張の特徴の把握するための情報整理手法の有効性を示した点で、政策過程の恣意性を排除した分析手法の確立に寄与するものと考えられる。「政策過程の恣意性」を排除するためには、利害関係者以外によって本稿のような分析がなされるであることがその一手段となる。しかしながら、そのみならず、冒頭で述べたように、これまで市民社会がその議事を逐一確認してその妥当性を検討することは極めて難しかった審議会等の議事について、議事録そのものではなく、クラスターと語群といった縮約した形で情報が示されることによって、議論の機会を増やし、なおかつ方法やテーマの名づけそれ自体の適切性も含めて議論することが可能になることも重要である。すなわち、本研究における分析手法の有効性の射程は政策過程研究にとどまらず、市民に分かりやすい可視化さ

れた形での公的議事録の公開，行政内部での効率的な審議会運営のためにも応用できるものであると考えられる。

なお，あくまで本論文の手法は「政策過程の恣意性」の排除を目指したものであり，「恣意性を排除した分析手法」とはなりえないことは，以上に指摘したものをはじめとした方法の選択や，ここでの手法におけるテーマの名づけが分析者の判断に依る以上，当然である。さらに，分析によって縮約された情報について，こうした分析上の恣意性までは排除できないことを踏まえて適切に解釈し，議論することが市民社会側にできなければ，単に議論の機会を増やすことに成功したとしても意味はない。

本稿の分析を実装するためには，公的議事録のフォーマットの標準化や，分析のアウトプットを可視化するツールの開発を合わせて進めていくことだけでなく，縮約した情報を示すことが市民社会における議論に対して持つ効果についてのさらなる検証が必要である。

参考文献

- 1) 西尾勝(2001)『行政学(新版)』有斐閣。
- 2) 樋口耕一(2004)「テキスト型データの計量的分析—2つのアプローチの峻別と統合」『理論と方法』19(1), 101-115.
- 3) 樋口耕一(2014)『社会調査のための計量テキスト分析』ナカニシヤ出版。
- 4) J.C.キャンベル・増山幹高(1994)「日本における診療報酬政策の展開」『季刊・社会保障研究』29(4), 359-368.
- 5) 池上直己・J.C.キャンベル(1996)『日本の医療—統制とバランス感覚』中央公論新社。
- 6) 結城康博(2004)『福祉社会における医療と政治—診療報酬をめぐる関係団体の動き』本の泉社。
- 7) 結城康博(2006)『医療の値段—診療報酬と政治』岩波書店。
- 8) 遠藤久夫(2014)「診療報酬の決定プロセス」『日本内科学会雑誌』103(12), 2892-2898.
- 9) 森田朗(2016)『会議の政治学Ⅲ』慈学社出版。
- 10) 広井良典(1994)「診療報酬の政治経済学—『自然増』をめぐる物語—」『医療と社会』3(2), 93-117.
- 11) 西岡晋(2001)「医療基本法案の政治過程」『早稲田政治公法研究』685: 191-222.
- 12) 内海和夫・乾孝司・橋本泰一・村上浩司・石川正道(2009)「テキストマイニングによる医療分野の課題及びその解決のための制度体系等に関する有用知識の抽出」『研究・技術計画学会・年次学術大会講演要旨集』24, 263-266.
- 13) 鄭蝦榮・小林潔司・羽鳥剛史(2009)「ファセット分解と公的討議のプロトコル分析」『土木計画学研究・講演集』

40(276).

- 14) 八村太輔・森幹彦・喜多一(2007)「議事録の構造化に基づくリフレクション支援」『電子情報通信学会技術研究報告・知能ソフトウェア工学』106(472), 49-54.
- 15) 難波雄二・塚井誠人・桑野将司(2010)「発言録データに基づく文脈マイニング手法の開発」『土木計画学研究・講演集』41(49).
- 16) 中谷隼・明村聖加・森口祐一(2014)「テキストマイニングによる東日本大震災前後の総合資源エネルギー調査会における論点の分析」『環境情報科学学術研究論文集』28, 113-118.
- 17) 岩見麻子・大野智彦・木村道徳(2011)「公共事業計画策定過程の議事録分析のための変動係数を用いた対象語選定手法の開発」『環境情報科学論文集』25, 55-60.
- 18) 岩見麻子・大野智彦・木村道徳・井手慎司(2012)「公共事業計画策定過程の議事録に対するテキストマイニングによる議論内容の把握に関する基礎的研究」『土木学会論文集 G(環境)』68(6), II_411-II_418.
- 19) 岩見麻子・大野智彦・木村道徳・井手慎司(2013)「公共事業計画策定過程の議事録分析によるサブテーマの把握とサブテーマを介した委員間の関係性の可視化に関する研究」『土木学会論文集 G(環境)』69(6), II_71-II_78.
- 20) 岩見麻子・大野智彦・木村道徳・井手慎司(2014)「公共事業計画策定過程の議事録分析による意見の協調・対立関係把握のための分析手法の開発」『土木学会論文集 G(環境)』70(6), II_249-II_256.
- 21) 安藤俊幸(2009)「テキストマイニングと統計解析言語 R による特許情報の可視化」『情報管理』52(1), 20-31.

謝辞

本研究は，科学技術イノベーション政策における「政策のための科学」推進事業により 2015 年 8 月に政策研究大学院大学に設置された科学技術イノベーション政策研究センター（GRIPS SciREX センター）のプロジェクト「政策形成のフレーミング、ステークホルダー分析、プロセスの構築を通じた政策形成プロセスの改善手法の開発」の下で実施したものであり，本研究成果は，森川の政策研究大学院大学の客員研究員としての成果を含んでいる。研究を支えてくださっている先生方ならびに職員の皆様，また貴重なコメントをくださった査読者の方々にこの場を借りて御礼申し上げます。

VISUALIZING STAKEHOLDERS' INTERESTS BY TEXT ANALYSIS: A CASE OF CENTRAL SOCIAL INSURANCE MEDICAL COUNCIL

Hiromu MATSUOKA¹, and So MORIKAWA²

¹ MPP (E-mail:hirom.matsuoka@gmail.com)

² Ph.D. Assistant Professor, The University of Tokyo, Dept. of Civil Engineering
(E-mail:morikawa@civil.t.u-tokyo.ac.jp)

This paper reports the result of our application of text analysis methods for the proceedings of Central Social Insurance Medical Council in Japan, aiming to visualize discussions in the Council. We look into applicability of several methods in the context of medical issues discussed in the Council and draw some implications for practical usage of the results of text analysis. We suggest that text analysis methods are partly successful in visualizing the discussion in two ways: (1) importance of the issues for each committee member can be captured by extracting themes consisting of several keywords by theme analysis especially by TFIDF method, and (2) unique features of committee members' arguments can be understood by further sub-theme analysis.

Key Words: *text analysis, visualization, Central Social Insurance Medical Council*