

新聞記事における政治家の発言の引用記述と 議会会議録との対応関係の調査 —フェイクニュース検出に向けて—

○^{1,2} 木村 泰知 ¹ 戸嶋 咲穂 ³ 渋谷 英潔

¹ 小樽商科大学 ² 理研 AIP ³ 横浜国立大学

kimura@res.otaru-uc.ac.jp

1 はじめに

近年、フェイクニュースが社会的な問題になっている [1]。フェイクニュースは、アメリカの大統領選にも影響を与えたと言われており、社会的混乱を防ぐためにも、自然言語処理技術による解決が期待されている [2]。自然言語処理関連の国際会議ではフェイクニュースに関する発表が増えており¹[3][4]、共通のデータセットを用いたフェイクニュースの shared task として「Fake News Challenge Stage 1 (FNC-I)」²が行われている。また、CLEF-2018 の「Fact checking Lab」³や NTCIR-14 の「QA Lab PoliInfo Task」⁴などでも政治情報の信憑性に関する shared task が設定されている。

情報の真偽を判断するために一次情報に当たることの重要性は広く知られており、例えば「〇〇議員が××と発言した」というニュースの真偽を判断するためには、〇〇議員の実際の過去の発言を調査すればよい。こういった議員発言の一次情報としては議会会議録などが考えられる。我々は、従来から、全国の自治体の地方議会会議録を収集・整理する手法を確立し、「地方議会会議録コーパス」の構築を進めてきた⁵[5][6]。新聞記事における引用箇所該当する発言を議会会議録などから探し、両者を比較可能な形で提示することができればフェイクニュースに対して有効であると考えられる。

そこで、本研究では、新聞記事で引用されている箇所をどのように探せばよいか検討するために、地方議

議会会議録コーパスを利用し、議会会議録の発言がどのように新聞記事に記述されているかを明らかにする。具体的には、豊洲問題に関する記事に対して、東京都議会会議録を対象として、引用箇所が含まれている発言箇所を見つける。豊洲問題は、2003年から議論されており、政治問題として新聞記事に記載されている。また、議会会議録は、いつ、どこで、だれが、なにを話したのかを記録したものであることから、新聞記事に引用されることがある。以上の理由から、豊洲問題と議会会議録を調査対象とした。例えば、下記の記事は、2017/3/30の日本経済新聞に「都議会閉会、豊洲移転持ち越し 自民は決断迫る」の見出しで掲載された3段落目である。この記事において引用箇所は、「多くの都民は移転問題の一刻も早い解決を求めている」および「知事は築地が直面している現実を直視するとともに、豊洲の市場としての適性を公平に判断すること」である。

日本経済新聞 2017/3/30

都議会閉会、豊洲移転持ち越し 自民は決断迫る
(2段落目まで省略)
最終日の討論で自民党は「多くの都民は移転問題の一刻も早い解決を求めている」と強調。築地市場の老朽化や土壌汚染の恐れについて問題提起した会期中の審議も踏まえて「知事は築地が直面している現実を直視するとともに、豊洲の市場としての適性を公平に判断すること」と迫った。

本論文では、上記のような新聞記事における発言の引用が、議会会議録に記載されている箇所を特定し、対応関係を調査した結果について述べる。

¹<http://www.aclweb.org/anthology/W/W17/#4200>

²<http://www.fakenewschallenge.org/>

³<http://alt.qcri.org/clef2018-factcheck/>

⁴<https://poliinfo.github.io/>

⁵<http://local-politics.jp>

2 豊洲問題と東京都議会

本論文では、新聞記事における議員発言の引用記述と議会会議録との対応関係の調査のために、「豊洲問題」の新聞記事を対象として、「東京都議会会議録」から引用箇所を見つける。まず、豊洲問題と東京都議会について述べる。

2.1 豊洲問題について

豊洲問題は、マスメディアでも取り上げられているように、築地市場の豊洲地区移転をめぐる問題である。ここで、簡単に豊洲問題について説明する。築地市場は、開場から80年以上経過しており、施設の老朽化、施設の搬入・搬出口や荷さばきのスペース不足、開放型の構造による衛生管理問題などから平成3年に築地の再整備が始まったが多くの問題が発生したため工事が中断された。平成15年には、移転整備へ方向転換し「豊洲新市場基本構想」が公表されている。平成16年には「豊洲新市場基本計画」が策定されている⁶。平成24年11月には、市場関係業界と豊洲の施設計画について合意している。しかしながら、土壌汚染問題の発覚などで「東京都中央卸売市場築地市場の移転・再整備に関する特別委員会」が設けられ証人喚問が行われている。平成30年10月11日の移転が決定している。

2.2 東京都議会

東京都議会は、年4回定期的（原則として2月・6月・9月・12月）に行われる「定例会」と必要に応じて行われる「臨時会」の2種類の議会がある。また、定例会や臨時会のはじめに会期が定められ、その期間中に「本会議」や「委員会」を開き議会活動が行われる。本会議は原則として議員定数の半数以上の議員が出席したときに行われる会議で、都議会に提出された議案や都議会としての意見表明等、議会の最終的な意思決定を行うものである。しかし、処理すべき分野は広範囲にわたり、限られた会期の中で多くの審議を行うためには本会議だけでは十分でないため、本会議の議決に先立って専門的かつ詳細に審査する機能を有するのが委員会である。現在、都議会委員会には条例によって設置された9つの常任委員会と、各会派の代表者等で構成される議会運営委員会、必要に応じて特定の事件を審査するため本会議の議決により臨時に設置される特別委員会がある。豊洲問題も、議決により特別委

員会が設けられている。そのため、本調査では「本会議」と「委員会」の全ての会議録を調査対象とする。

3 引用記述の調査

3.1 調査の目的

本調査では、豊洲問題を対象として、新聞記事における発言の引用記述が、東京都議会会議録において、どのように記載されているかを明らかにする。

3.2 分析方法

分析は、下記の手順で進める。

1. 豊洲問題に関する記事の収集
2. 東京都議会における引用箇所の特定
3. 新聞記事の引用箇所と都議会会議録の発言箇所の対応表作成

3.2.1 新聞記事のデータ

豊洲問題に関する新聞記事の「対象とする新聞」「検索条件」「対象期間」は次の通りである。

- 対象：日本経済新聞電子版の記事
- 条件：キーワード「豊洲」「答弁」をAND検索した結果、該当する記事
- 期間：2003年3月～2017年9月

調査対象の新聞は「日本経済新聞電子版」（以降、日経電子版とする）とする。対象記事は、「豊洲」「答弁」をAND検索した結果とする。検索条件の「豊洲」というクエリにより対象を絞り込み、「答弁」という議会特有の表現を利用することで、政治問題の記事に絞り込む。「答弁」とは「質問に対して答えること。またはその答え」を示す単語であり、「知事が答弁した」「議会で答弁する」のように使われており、議会特有の語である。対象期間は、豊洲問題が激化したのは平成28年であるが、豊洲新市場基本構想が策定されたのは平成15年のことであるため、調査期間を平成15年の3月から平成29年の9月までとする。

⁶http://www.shijou.metro.tokyo.jp/toyosu/siryou/keikaku_siryou/

分類	文数	会議録	記事	
(1)	34文	「既定路線が敷かれている中で、小池知事が待ったをかけたことは、当時、大変勇気の要る決断であったと思います。」	「既定路線が敷かれている中で勇気の要る決断」	
(2)	(2a)	34文	「この機会に東京の施策や魅力を広くPRするとともに、世界の主要都市のリーダーたちとも大いに意見交換をしてみたい。」	「この機会に東京の施策や魅力を広くPRするとともに世界の主要都市のリーダーとも意見交換してくる」
	(2b)	39文	「議会の意思については、これを尊重していきたいと思っております。」	「議会の意思を尊重する」
	(2c)	35文	「顧問政治または側近政治を重視することは、都民ファーストを標榜する知事の目指す方向とは逆行するのではないのでしょうか。」	「『側近政治』は都民ファーストを標榜する知事の姿勢と逆行する」
(3)	8文	「もし、小池知事が開場に待ったをかけることなく、当初予定のまま、昨年十一月七日に新市場が開場していた場合には、こと一月の基準値を超える地下水モニタリング結果だけでなく、新市場営業中に盛り土がなかった事実が判明した可能性もあります。その場合には、豊洲新市場の風評被害の懸念は現実のものとなって、豊洲新市場は開場当初から大混乱に陥っていたものと考えられます。」	「（豊洲開場後に土壌汚染対策の不備が見つかったら）大混乱に陥っていた」	

図 1: 新聞記事の引用箇所と都議会会議録の発言箇所の対応例

3.3 調査結果

表 1 は、日経電子版において「豊洲」と「答弁」の両方を含む記事 67 件のうち、引用箇所がある 32 記事であり、掲載日、見出し、対応関係の文数、記事の文字数、引用箇所の文字数である。引用箇所は、67 記事のうち 32 記事であり、対応関係の文数は 150 文であった。掲載期間は、豊洲新市場基本構想が策定された 2003 年以降を対象としたが、最も古い記事が 2010 年であったため、収集期間が 2010 年 3 月から 2017 年 9 月までとなった。

対応箇所を自動的に推定する難しさの観点から、32 記事における 150 文を以下のように分類した。

- (1) 新聞記事の引用箇所と都議会会議録の発言箇所の表現が完全に一致する。
- (2) 新聞記事の引用箇所と都議会会議録の発言箇所の表現がほぼ一致する。
 - (2a) 一致しない箇所は、敬語を常態語（普通語）に直していたり、敬体を常体に直している。
 - (2b) 一致しない箇所は、助詞、接続詞、形式名詞などの実質的な意味を表さない語句を言い換えている。
 - (2c) (2a) と (2b) に当てはまらない、実質的な意味を表す語句を言い換えている。
- (3) 新聞記事の引用箇所と都議会会議録の発言箇所の表現が大きく異なる。

まず、新聞記事の引用箇所と都議会会議録の発言箇所の一致度合から 3 つ（1-3）に分類し、基本的に一致するが一部が一致しない 2. に関しては、その一致しない箇所に応じてさらに 3 つ（2a-2c）に分類した。

図 1 は、上記の分類に該当する、新聞記事の引用箇所と都議会会議録の発言箇所の対応関係の例を挙げている。図中の太字は対応関係にある箇所を示し、下線部は表現的に一致しない箇所を示している。(1) は、語句が連続しているかどうかの問題はあるが、引用箇所に該当する表現が全て発言箇所に存在しており、exact match により推定できると考えられる。(1) は最も簡単に対応関係を推定できる場合であり、調査対象の 22.7% (34/150 文) を占めた。(2) は、引用箇所を Bag of Words (BoW) として表現することで対応関係が推定できそうなものである。特に (2a) と (2b) に関しては、助詞や形式名詞などの機能語を除いて BoW を作成することで問題なく推定できると考えられる。一方で、(2c) は「目指す方向」と「姿勢」のように、類義語または類義表現を考慮する必要があり、(2a) や (2b) と比較して推定が難しい。(2a), (2b), (2c) が占める割合は、それぞれ、22.7% (34/150 文), 26.0% (39/150 文), 23.3% (35/150 文) であった。(3) は、対応関係にある箇所全体の意味を考慮しなくてはならず、語句レベルの対応関係を超過しているため最も難しい推定となる。しかしながら、(3) は調査対象の 5.3% (8/150 文) と少なく、引用箇所の約 95% は BoW などの語句レベルの一致により推定できることが分かった。

4 おわりに

本論文では、豊洲問題を対象として、日経新聞記事における発言の引用が、議会会議録に記載されている箇所を特定し、対応関係を調査した結果について述べた。調査分析では日経電子版の記事から「豊洲問題」に関する記事 67 点を収集し、そのうち発言文（会議録要約文）が載る 32 点の記事から 150 の対応関係を抽出し、分析を行った。

今後、我々は、NTCIR の QA Lab 4 Poliinfo にお

表 1: 「豊洲」と「答弁」を含む記事のうち、引用箇所がある 32 記事の情報

記事番号	掲載日	見出し	対応関係 の文数	記事の 文字数	引用箇所 発言文字数
1	2010/3/16	築地市場移転、都議会民主が予算修正案提出を検討	4	1078	103
2	2010/3/27	築地市場移転、予算案可決へ 予定地購入は先送り	4	591	95
3	2010/3/31	築地移転、結論出ず 都議会、10 年度予算案可決	1	809	10
4	2010/12/8	東京都、水道事業の海外進出後押し 企業に情報提供	4	501	95
5	2011/2/16	石原知事と民主、都議会で溝深く 新銀行・築地移転で論戦	4	481	113
6	2011/6/24	都、東北 3 県への旅行に 1 泊 3 千円助成 5 万泊分用意へ	1	488	18
7	2016/9/22	豊洲市場の盛り土 いつ、誰が、なぜ変更? 小池都知事が追加調査指示	1	878	21
8	2016/9/27	市場長、都議会で謝罪 豊洲盛り土問題	4	516	128
9	2016/10/4	豊洲問題「トップとして重く受け止める」 小池都知事	8	560	230
10	2016/10/4	都議「知事こそブラックボックス」 衝突は回避都議会、静かな幕開け	12	1072	264
11	2016/10/4	小池知事「五輪は 1 カ月メドに結論」 都議会答弁	6	749	125
12	2016/10/5	頼もしい「不十分だ」小池知事答弁に評価と不満	1	1049	29
13	2016/10/5	小池都知事「責任者は懲戒処分」 豊洲問題	7	469	200
14	2016/10/5	小池知事、歴代幹部の処分検討「体制を刷新」	4	632	88
15	2016/10/6	豊洲盛り土問題「知らなかった」 都議会集中審議	7	762	180
16	2016/10/8	豊洲報告の誤り「責任転嫁だ」 技術会議元委員ら、都に怒り	2	839	62
17	2016/10/13	小池都知事、市場長を更迭へ 豊洲盛り土問題で体制刷新	1	378	11
18	2016/10/13	小池色、カギ握る政策 初の都議会 静かに閉会「金融都市」「AI」肉付け必要	6	1504	185
19	2016/10/14	小池知事、市場幹部 10 人規模で刷新 14 日発表へ市場長更迭も	1	976	11
20	2016/10/31	小池都政 3 カ月、急進改革は破壊か創造か予算が試金石	2	1706	65
21	2016/12/6	15 年度中央卸売市場会計を不認定 都議会特別委	1	370	50
22	2016/12/8	都議会自民、一転対決姿勢に 知事との事前質疑調整なし	7	1061	183
23	2016/12/15	批判の応酬 審議迷走 都議会閉会、知事・自民対立鋭く	6	1292	108
24	2017/3/11	豊洲移転「自分が進言、石原氏も『早く』と指示」 元市場長	8	1075	190
25	2017/3/19	「都の体質浮き彫り」 都議や傍聴者ら不満の声豊洲百条委	1	462	19
26	2017/3/21	石原氏追及は空振り 百条委で「報告ない」「一任」連発	10	953	172
27	2017/3/30	都議会閉会、豊洲移転持ち越し 自民は決断迫る	2	1100	69
28	2017/6/3	市場移転判断、明言せず 都議会で小池知事答弁	8	469	187
29	2017/8/26	豊洲論戦、新議会で幕開け 経済・港湾委で質疑	3	902	78
30	2017/8/31	豊洲移転 延期表明 1 年 都政透明化 光と影	10	1139	220
31	2017/8/31	蜜月、追及…小池知事と初の論戦 都民フ、賛同に終始	8	1188	197
32	2017/9/27	小池新党旗揚げ 都議会公明「都政を踏み台」	6	954	167
			150	27003	3673

いて、入力を新聞記事として、出力を地方議会会議録に含まれる関連発言とするタスクも検討している。

謝辞

本研究は JSPS 科研費 JP16H02912 の助成を受けたものです。

参考文献

- [1] BuzzFeedNews.ThisAnalysisShows HowViral-FakeElectionNewsStories OutperformedReal-NewsOnFacebook, November 17, 2016.
- [2] Wang, William Yang, "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection , Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 422–426, 2017.
- [3] Thorne, James et al., Fake news stance detection using stacked ensemble of classifiers, Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism, pp. 80–83, 2017.
- [4] Bourgonje, Peter et al., From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles, Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism, pp. 84–89, 2017.
- [5] 筒井貴士, 我満拓弥, 大城卓, 菅原晃平, 永井隆広, 渋谷英潔, 木村泰知, 森辰則, 地方議会会議録コーパスの構築および政治情報システム構築を目標としたアノテーションの一提案. 自然言語処理, Vol.21, No.2, pp.125-156,2014.
- [6] Yasutomo Kimura et al., Creating Japanese Political Corpus from Local Assembly Minutes of 47 Prefectures, Coling 2016, The 12th Workshop on Asian Language Resources, pp.78-85, 2016.