

NTCIR-15 QA Lab-PoliInfo-2 における Dialog Summarization

小川 泰弘[†] 木村 泰知[‡] 渋木 英潔[§] 秋葉 友良[¶] 横手 健一^{*} 乙武 北斗[◇] 石下 円香[§]

[†] 名古屋大学 [‡] 小樽商科大学 [§] 国立情報学研究所

[¶] 豊橋技術科学大学 ^{*} 日立製作所 [◇] 福岡大学

[†] yasuihiro@is.nagoya-u.ac.jp

[‡] kimura@res.otaru-uc.ac.jp

1 はじめに

近年、政治にまつわるフェイクニュースが社会的な問題とされてきている。信憑性の低い情報がソーシャルメディアを介して拡散され、民意の形成に偏りを生じさせることが懸念されている。また、政治家の発言自体も信憑性や根拠が曖昧な場合が多く、政治家の発言に対するファクトチェックの必要性も高まっているといえる。

近年、EuroParl Corpus や UK Hansard コーパスなどが、政治に関わる自然言語処理研究のための言語資源として利用されている。しかしながら現在のところ、日本語を対象とした研究データが少ないことに加えて、議会における議員の発言を対象とした研究や、議会における政策形成のための議論を対象とした研究は進んでいない。

これらの問題を解決することを目指して、評価型ワークショップNTCIRにおいてShared task QA Lab-PoliInfoを行っている。2019年後半から2020年にかけて実施しているNTCIR-15 QA Lab-PoliInfo-2は三つのサブタスクで構成される。Stance Classification サブタスクでは、議会における発言に基づいて会派の立場（議案への賛否）を抽出することを目指す。Dialog Summarization サブタスクでは、本会議で行われる一般質問および代表質問における質疑の自動要約を目指す。Entity Linking サブタスクでは、発言内で言及されている法律名を抽出し、知識ベース・言語資源の該当項目と結びつけることを目指す。

それらのうち、本稿においては、Dialog Summarization サブタスクの紹介と、実際にサブタスクに取り組んだ一例を紹介する。

本稿では、まず第2節でNTCIRにおけるQA Labタスクの背景について述べた上で、第3節でQA Lab-

PoliInfo-2 タスクにおける Dialog Summarization の概要と、タスクのためのデータセットの構築について述べる。第4節では、タスクオーガナイザがサブタスクに取り組んだ結果について紹介する。

2 QA Lab-PoliInfo タスクの背景

QA Lab-PoliInfo は、政治情報を対象としたサブタスクを質問応答の枠組みで捉えたものとして、評価型ワークショップNTCIR¹の第14回において開催した。地方議会会議録コーパス [4] を対象に、議会発言における議員の意見やその根拠、立場などを抽出し、議論の構造を明確にして、関係性などを理解しやすいように整理して提示することを最終的な目標としている。

NTCIR-14 においては、Segmentation, Summarization, Classification の三つのサブタスクを実施した。Segmentation は、要約されて引用された議会中の発言に対して、会議録の引用元の範囲を特定することを目指したサブタスクである。Summarization は、会議録中の特定の範囲から、発言者の意図が誤解されないように指定された文字数の範囲内で要約文を生成するサブタスクである。Classification は、議会における発言に、政治課題に対する意見が含まれるか、また意見に対する事実検証可能な根拠が示されているかを分類するサブタスクである。QA Lab-PoliInfo タスクには、15組のチームが参加し、2018年11月から12月にかけて Formal Run が行われた [2][3]。

本稿で述べる NTCIR-15 QA Lab-PoliInfo-2 は、NTCIR14 QA Lab-PoliInfo の成果と問題点を踏まえて、Stance Classification, Dialog Summarization, Entity Linking の三つのサブタスクを提案している。

本タスクは、ウェブサイト²を通じて、参加者に共

¹<http://research.nii.ac.jp/ntcir/index-ja.html>

²<https://poliinfo2.net/>

通のデータセットを配布するとともに、評価結果を随時表示するリーダーボードを設置することで、政治情報を対象とした自然言語処理研究の促進を目指す。

3 Dialog Summarization

本節では、Dialog Summarization サブタスクの概要と、構築したデータセットの設計について述べる。

3.1 タスクの概要

政治家の発言の信憑性を判断するためには、政治課題に関する議論がどのように行われているのか、知る必要があり、議論をしている相手の発言や文脈を考慮しなければならない。政治課題に関する議論は、議会において行われており、議会会議録として質問や答弁が残されている。しかしながら、議会会議録は発言を書き起こした文書であり、まとめられておらず読みづらいという問題がある。特に、東京都議会をはじめとする多くの地方議会では、一問一答方式ではなく、一括質問一括答弁方式がとられており、質問と答弁が離れた位置に存在する。また、質問に対して、知事が答弁する場合と、総務部長や教育長のような知事以外の出席者が答弁する場合がある。さらには、知事による答弁を補足する形で複数の出席者が答弁することもある。よって、質疑（質問と答弁の組）を要約するためには、議論の構造を考慮することが求められる。以上のことから、Dialog Summarization は地方議会における「議員の質問」と「知事側の答弁」という対話構造を考慮しながら要約することを目標としている。

本サブタスクの入力は東京都議会の会議録、および、都議会だよりの要約に必要な情報であり、出力は都議会だよりの形式に沿った要約結果である。評価について、複数の手法を並行して用いる予定であるが、公式サイトリーダーボードでは、ROUGE-1 Recall を用いて順位を決める。

3.2 データセットの構築

議会における一般質問および代表質問の概要は都議会だよりに掲載されている（図1右）。都議会だよりは、議会に記載された内容が議会事務局の職員により作られていることから、人手により作成された「正解の要約」とみなすことができる。また、都議会だよりで、質問項目ごとに質問者と答弁者が示されており、質問とそれに対応する答弁が簡潔にまとめられている。そこで、都議会だよりに記された質疑の要約を Dialog Summarization の正解として用いることとする。図1

表 1: Dialog Summarization のデータ構造

ID	識別番号
Date	日付
Prefecture	都道府県
Meeting	会議名
MainTopic	メイントピック
QuestionSpeaker	質問者
SubTopic	サブトピック
QuestionSummary	質問の要約
QuestionLength	質問の字数制限
QuestionStartingLine	質問の開始行
QuestionEndingLine	質問の終了行
AnswerSpeaker	答弁者 ※リスト型
AnswerSummary	答弁の要約 ※リスト型
AnswerLength	答弁の字数制限 ※リスト型
AnswerStartingLine	答弁の開始行 ※リスト型
AnswerEndingLine	答弁の終了行 ※リスト型

の例では、左側が要約前の議会会議録であり、右側に黄色で示した範囲の文字列が正解データとなる。

Dialog Summarization サブタスクでは、東京都議会会議録に加えて、機械処理が容易な JSON 形式で、要約に必要な情報を含んだ出題ファイルを提供する。出題ファイルのデータ構造を表1に、JSON 形式のデータの例を図2にそれぞれ示す。

出題ファイルでは、識別番号 (ID)、議会の開催日 (Date)、都道府県名 (Prefecture)、会議名 (例えば「平成 31 年度第 1 回定例会」) (Meeting)、議会だよりに書かれている一人の質問者の全質問を端的に表す単文 (図1の例では「中小企業・小規模企業の支援を幼児教育無償化への都の対応は」) (MainTopic)、質問項目 (図1の例では「産業振興」や「ダイバーシティ・東京」) (SubTopic)、質問者名および答弁者名 (Question-Speaker, AnswerSpeaker)、質問および答弁の開始行・終了行 (QuestionStartingLine, QuestionEndingLine, AnswerStartingLine, AnswerEndingLine)、生成する質問・答弁の要約の字数制限 (QuestionLength, AnswerLength) が与えられる。これに対して参加者は会議録に基づいて、質問と答弁の要約を生成し、QuestionSummary, AnswerSummary にそれぞれの解答を出力する。

4 タスクオーガナイザによる実践

Dialog Summarization サブタスクにおいては、タスクオーガナイザによる実行結果と、その際に作成したプログラムやモデルなどを提供予定である。本節では、タスクオーガナイザによる実践について述べる。

入力（都議会会議録）

百十五番小山川くひこ君
(百十五番小山川くひこ君登壇)

○百十五番 小山川くひこ君 東京都議会第四回定例会に当たり、都民ファーストの会東京都議団を代表して、小池知事及び教育長、関係局長に質問いたします。

いよいよ二〇二〇年の東京オリンピック・パラリンピック競技大会まで二年を切りました。一九六四年の東京大会は、戦後復興の象徴であり、首都高速道路や地下鉄の建設、東海道新幹線の開通など、各種インフラの整備が進みました。一九六四年大会後、日本は高度経済成長を続け、その後の日本と東京の発展へと大きくつながりました。

その後の平成は激動の時代でありました。バブル崩壊から始まった長期的な経済停滞、経済のグローバル化、IT化の流れの中で、日本の国際的地位は低下しました。一九六四年大会後に増加を続けていた日本の人口は、二〇〇八年をピークに減りに減少に転じており、東京都の人口も二〇二五年をピークに減少に転じると見込まれております。このような社会経済情勢の劇的な変化は、戦後日本の成長を生んだ社会モデルからの変革を迫っております。

平成の時代が閉幕を閉じ、新たな時代を迎える成熟都市東京は、今まさに大きな変革を必要としています。少子高齢化による生産年齢人口が減少する中で、次なる成長の源泉となる人、物、金、情報をめぐる世界の都市間競争、まさに機軸をきわめていま、このような状況下において多様なご成長の源泉である。そういった認識に立ち、二〇二〇年の東京大会とご先を契機として、世界の中で競争東京の成長戦略を遂行させなければなりません。

そして、私たちは、一九六四年東京大会をきっかけに築き上げられてきた東京を二〇二〇年大会を契機として再構築し、東京と他の地域がともに栄える、東京の持続的成長を実現していかなければなりません。

私たち都民ファーストの会東京都議団は、都議会最大会派となり一年余り経過しました。この間、議会改革を初め、受動喫煙防止条例の制定、待機児童の大削減、オリンピック・パラリンピック憲章人権条例の成立など、二〇二〇年の先を見据えた東京の成長と発展の礎となる施策が着実に推進されてきました。

本定例会でも、中小企業への振興策、防災対策、働き方改革を柱とする補正予算など、未来の東京の成長と発展のために必要不可欠な施策を取り上げております。

このような東京の取り組みにもかかわらず、国はまた、不合理な都税の収奪を繰り返そうとしています。今、都議会を求められているのは、都議会一丸となって、他の地域と共存共栄可能とする首都東京の成長戦略を描き出し、着実に実行することであると改めて申し上げ、以下質問いたします。

平成三十一年度税制改正について伺います。

国は、いわゆる偏在正定の名のもと、都の税財源を地方へと配分すべく、さまざまな措置を講じてきました。この間、都としても対応策を講じてきましたが、平成に入ってから三十年間で都が失った財源は六兆円以上、平成三十一年度税制改正においても、さらなる措置が事実上予定されております。

こうした国の不合理な税制改正の動きに対して、先般、私たちの提案により立ち上げられました東京と日本の成長を考える検討会の報告書が取りまとめられ、また、東京都税制調査会の答申も示されました。そして、それらを受けた東京都の見解も示されております。

都はこれまで、小池知事を先頭に、全国知事会や東京都選出の国会議員、与党税制調査会の国会議員、都内区市町村との協働を行ってまいりました。私たち都民ファーストの会東京都議団も、東京都選出の国会議員や与党税制調査会の国会議員への要請活動、都民への啓発活動等に努めてまいりました。

出力（都議会だより）

東京都議会
Tokyo Metropolitan Assembly

都議会の紹介 議員の紹介 会議の結果と記録 傍聴・見学 調査・友好交流など

トップ > 都議会だより > 331号 > 小池知事 小山川くひこ君

中小企業・小規模企業の支援を 幼児教育無償化への都の対応は

小山川くひこ (都民ファースト)

産業振興

(1) 中小企業・小規模企業振興条例の理念に基づき、活力ある地域社会をつくり創出の推進。(2) 農業は東京の持続的成長に必要不可欠。農業振興への今後の取組は、

知事 (1) 地域経済の持続的発展と雇用創出の実現のため効果の高い施策を展開。(2) 都市農地の保全、担い手の確保と育成・定着の体制整備、先進技術活用等、様々な施策を展開。

ダイバーシティ・東京

(1) 幼児教育無償化では負担の軽減は十分とは言えず、また認可外認可外で格差が生じる。対応は、(2) 児童虐待対策の基幹制度では発覚防止の強化を重視して進める。(3) LINE 相談の一種の活用を含め発覚は、(3) 小中学校のスクール・サポート・スタッフの配置支援を拡大す。 (4) 学校の働き方改革を加速させるため、部活動指導員をはじめ専門スタッフの質・量の確保を。(5) 受動喫煙防止条例の施行に向けて、内容の周知徹底と実効性の確保を。

知事 (1) 待機児童対策協議会で国と意見交換。国の動きを踏まえ適切に対応。(2) 体罰等を行ってはいけないこと等を未然防止の観点から条例に明記。LINE相談は31年度から本格実施。(5) 条例施行等のタイミングで効果的な広報を展開。都民や事業者の理解促進や機運の醸成を図り、受動喫煙防止の取組を進める。

教育長 (3) 区市町村教育委員会と連携しながら配置拡充を図る。(4) スタッフの安定的確保や賃金向上をはじめとする多様な取組を検討。

図 1: Dialog Summarization のデータセット構築方法

```

1 { "AnswerEndingLine": [540],
2   "AnswerLength": [50],
3   "AnswerSpeaker": ["知事"],
4   "AnswerStartingLine": [533],
5   "AnswerSummary": ["新視点の対応加えた防災対応指針を 11
6     月に策定し防災力向上の道筋を早期に示す。"],
7   "Date": "2011-06-23",
8   "ID": "PoliInfo2-DialogSummarization-JA-Dry-
9     Training-Segmented-00002",
10  "MainTopic": "東京の総合防災力を更に高めよ<br>環境に配慮
11    した都市づくりを",
12  "Meeting": "平成 23年第 2回定例会",
13  "Prefecture": "東京都",
14  "QuestionEndingLine": 291,
15  "QuestionLength": 50,
16  "QuestionSpeaker": "山下太郎(民主党)",
17  "QuestionStartingLine": 284,
18  "QuestionSummary": "更に高める取組が必要。",
19  "SubTopic": "東京の総合防災力"}

```

図 2: Dialog Summarization のデータサンプル

4.1 前処理

要約元のテキストが含まれる議会会議録のファイルと、正解要約となる議会だよりには細かな差異が存在するため、前処理が必要となる。例えば、議会会議録においては「石原慎太郎」のように発言者の氏名が記述されているが、議会だよりにおいては答弁者は「知事」のように役職名で記述されている。また、質問者ごとの質疑全体の範囲を特定する必要がある。こうした前処理については、全参加者共通であるため、対処するためのデータをタスクオーガナイザ側で提供する。

4.2 会議録の構造解析

第3節で述べたように、本サブタスクにおいては、まず会議録の構造を解析し、サブピックごとに要約元となる質問や答弁の範囲を推定する必要がある。これは

NTCIR-14 QA Lab-PoliInfo における Segmentation サブタスクと類似したタスクである。しかし、先の Segmentation サブタスクにおいては、要約後のテキストを利用して要約元の範囲を推定することができたが、今回のサブタスクにおいては、要約後のテキストは利用できない。しかし、要約後のテキストに付加されたサブトピックや発言者の情報は利用可能である。また、先の Segmentation サブタスクにおいて最も優れた成果を達成した文献 [1] においては、表 2 に示す手掛かり句に基づいてテキストを分割することが有効であることが示されている。

よって、今回は表 2 の手掛かり句に基づいてテキストを段落に分割後、サブトピックを含む段落を要約元の候補とした。なお、サブトピックが「築地市場移転問題」のように長い場合、「築地市場」や「移転」などの用語は会議録に出現しても、「築地市場移転問題」自体が出現しない場合がある。そうした場合は BERT を用いて文およびサブトピックをベクトル化し、最も類似する文を含む段落を候補とした。その後、議会だよりにおけるサブピックの出現順と、会議録における要約元の出現順が同じであるという仮定を利用したヒューリスティクスにより、要約元の範囲を決定した。

4.3 要約

テキストの要約には、NTCIR-14 QA Lab-PoliInfo の Summarization サブタスクにおける自動評価において、最も良い成果を示した文献 [5] の手法をそのまま利用した。文献 [5] では、正例と負例の比率を変化

表 2: 手掛かり句を見つけるための正規表現

pattern	regular expression
opening	^まず ^最初に ^初めに ^次に ^次いで ^最後に ^終わりに ^-[一二三四五六七八九十]+点目 ^-[、,]+についてで(す あります ございます)(が けれど) ^-終わり(ま で)す。 ^以上で ^ありがとうございます 他の質問に(ついて つきまして)は
closing	伺い [^、]*ます。 お尋ね [^、]*します お答えください。 (見解 所見 答弁)を求め [^、]*ます。 (いかがで どうで)(しょうか すか)。 .+質問を(終わります 終了します)。

(文献 [1] からの引用)

表 3: ROUGE による評価結果

	Recall							F-measure						
	N1	N2	N3	N4	L	SU4	W1.2	N1	N2	N3	N4	L	SU4	W1.2
表層形	0.404	0.130	0.076	0.046	0.337	0.168	0.160	0.307	0.098	0.058	0.035	0.253	0.123	0.159
原形	0.424	0.143	0.086	0.054	0.354	0.184	0.170	0.323	0.108	0.065	0.041	0.266	0.134	0.169
内容語	0.244	0.105	0.050	0.022	0.234	0.123	0.150	0.185	0.078	0.037	0.017	0.177	0.079	0.139

させた訓練データから学習した複数のランダムフォレストを組み合わせて要約する。先の Summarization サブタスクと今回の Dialog Summarization サブタスクでは、用意された訓練データの量が異なる。先の Summarization サブタスクでは平成 24 年分のみであったが、今回 Dialog Summarization サブタスクでは、それに加えて平成 23 年分、平成 26 年分が利用可能である³。しかし、本稿では追加の訓練データは利用せず、先の Summarization サブタスクで学習したモデルをそのまま利用した。その結果を表 3 に示す。

5 おわりに

本稿では、NTCIR-15 QA Lab-PoliInfo-2 における Dialog Summarization タスクの概要、データセット、およびタスクオーガナイザによる実践について述べた。

NTCIR-15 QA Lab-PoliInfo-2 は、2020 年 2 月 1 日から 4 月 30 日まで参加を受け付けている。本稿で述べたデータセットを用い、参加者相互で議論を行いながら、5~6 月に予備テスト (Dry Run) を、7 月に本テスト (Formal Run) を実施する予定である。

公式サイトでは、参加者の結果を随時掲載するだけでなく、提出データを共有し、例えば、他の参加者の構造解析結果を利用した要約なども可能とする。

³テストデータは平成 25 年度分である。

参考文献

- [1] Kanasaki K, Yong J, Kawamura S, et al. Cuephrase-based text segmentation and optimal segment concatenation for the NTCIR-14 QA Lab-PoliInfo task. In *NII Testbeds and Community for Information Access Research - 14th International Conference, NTCIR 2019, Tokyo, Japan, June 10-13, 2019, Revised Selected Papers*, pp. 85–96, 2019.
- [2] Kimura Y, Shibuki H, Ototake H, et al. Final report of the NTCIR-14 QA lab-poliinfo task. In *NII Testbeds and Community for Information Access Research - 14th International Conference, NTCIR 2019, Tokyo, Japan, June 10-13, 2019, Revised Selected Papers*, pp. 122–135, 2019.
- [3] Kimura Y, Shibuki H, Ototake H, et al. Overview of the NTCIR-14 QA Lab-PoliInfo task. In *Proceedings of the 14th NTCIR Conference*, 2019.
- [4] Kimura Y, Takamaru K, Tanaka T, et al. Creating Japanese political corpus from local assembly minutes of 47 prefectures. In *Proceedings of the 12th Workshop on Asian Language Resources (ALR12)*, pp. 78–85, 2016.
- [5] Ogawa Y, Satou M, Komamizu T, et al. nagoy team's summarization system at the NTCIR-14 QA Lab-Poliinfo. In *NII Testbeds and Community for Information Access Research - 14th International Conference, NTCIR 2019, Tokyo, Japan, June 10-13, 2019, Revised Selected Papers*, pp. 110–121, 2019.