

NTCIR-14 QA Lab-PoliInfo の Formal Run Dataset の構築

木村泰知^{1,2} 渋木英潔³ 乙武北斗⁴ 内田ゆず⁵ 高丸圭一⁶
 阪本浩太郎^{3,7} 石下円香⁷ 三田村照子⁸ 神門典子^{7,9}

¹ 小樽商科大学 ² 理化学研究所 AIP ³ 横浜国立大学 ⁴ 福岡大学 ⁵ 北海学園大学
⁶ 宇都宮共和大学 ⁷ 国立情報学研究所 ⁸ カーネギーメロン大学 ⁹ 総合研究大学院大学

kimura@res.otaru-uc.ac.jp

1 はじめに

QA Lab タスクは、評価型ワークショップ NTCIR において、現実世界における質問応答システムの実現を目指して開催されている。NTCIR-11 の QA Lab[1] を第 1 回として NTCIR-13 の QA Lab-3 まで、世界史の大学入試問題を対象として実施してきたが、4 回目の QA Lab-PoliInfo では政治情報を対象とした質問応答に取り組んでいる。

QA Lab-PoliInfo では、地方議会会議録を用いて、議員の発言に含まれる意見やその根拠や条件などを抽出し、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。

NTCIR14 QALab-PoliInfo Formal Run は、2018 年 11 月から 12 月にかけて開催された。本稿では、Formal run の 3 つのタスクの意義と Formal Run Dataset¹ の構築方法について述べる。

2 NTCIR14 QALab-PoliInfo

QALab-PoliInfo では、政治情報の信憑性を判断する処理を関連記述抽出、重要記述抽出、信憑性判断支援要約の三段階に分けている。そして、それぞれの要素技術として、Segmentation Task, Classification Task, Summarization Task を位置付けている。図 1 に QALab-PoliInfo の 3 つのタスクのイメージを示す。以下、説明の都合上、Segmentation Task, Summarization Task, Classification Task の順で説明をする。

¹Formal Run Dataset は 2019 年 3 月から、<https://github.com/kmr-y/NTCIR14-QALab-PoliInfo-FormalRunDataset> において公開する。

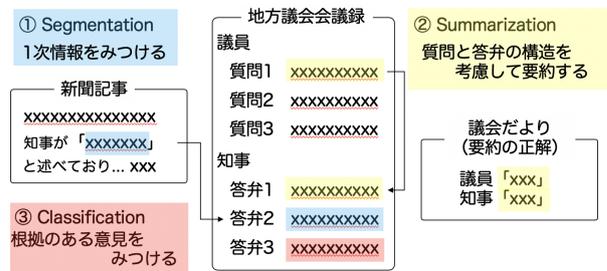


図 1: QALab-PoliInfo の 3 つのタスクのイメージ

2.1 Segmentation Task

Segmentation task は、新聞記事やマイクロブログなどで引用されている二次情報に対して、発言したとされる一次情報の該当範囲を提示することを目的としている。

本タスクの入力、出力、評価は、次の通りである。

入力: 質問と答弁の「発言の引用の対 (ついで)」とその発言が含まれる「議会会議録」
 出力: 引用を理解するために読むべき発言の範囲 (「開始行」「終了行」)
 評価: 開始行と終了行の一致率

表 1 にデータ形式を示す。図 2 の例では、質問の「[1] 区市町村と連携を。[2] 国にナショナルセンター設置を強く求めるべき。」と答弁の「[1] 研修会等で支援センター活用を働きかける。[2] 支援策充実を働きかける。」という引用に対して、質問は 23,037 文目から 23,048 文目まで、答弁は 23,234 文目から 23,239 文目までが該当するという出力が求められている。

QA Lab-PoliInfo では「都議会だより」の記事を引用文 (Summarization Task では参照要約) とみなし、議会会議録の発言と対応付けることで正解データを作成した。東京都をはじめ多くの自治体では、議会の広報や情報公開として、「都議会だより²」のように質問と答弁を簡潔にまとめたものを公開している。データ作成では、二人の注釈者が「都議会だより」に記載されている内容を引用された二次情報と仮定して、都議会会議録の発言を一次情報として、適切な範囲をみつける。二人の注釈が一致した場合はそのまま利用し、一致しない場合は、二人に協議してもらい一致させた。表 2 に Training data 数と Test data 数を示す。

表 1: Segmentation Task のデータ形式

表記	説明
ID	識別子
Prefecture	都道府県名
Date	日付
Meeting	会議名 (議会だよりの記載)
MainTopic	メインピック
SubTopic	サブピック
QuestionSpeaker	質問者
QuestionSummary	質問の要約
AnswerSpeaker	答弁者
AnswerSummary	答弁の要約
QuestionStartingLine	質問の開始行
QuestionEndingLine	質問の終了行
AnswerStartingLine	答弁の開始行
AnswerEndingLine	答弁の終了行

表 2: Segmentation のデータ数

Training data	Test data
298	83

2.2 Summarization Task

Summarization Task は、議会の質問と答弁の構造を考慮して「発話者の意図を歪めない要約」を生成することを目的としている。

本タスクの入力、出力、評価は、次の通りである。

入力: 議会会議録に含まれる「発言」と要約の「制限字数」と「議会会議録」
 出力: 発話者の意図が伝わる「要約」
 評価: ROUGE, 参加者による評価

表 3 にデータ形式を示す。図 3 は Formal run の json 形式の例を示す。Summarization Task では、ある議員の議会会議録中の「発言」と要約の「制限文字数」が与えられ、要約を作成する。

データ作成は、Segmentation task の作成方法と同様である。表 4 に Training data 数と Test data 数を示す。Test data は Multi と Single の 2 つのタイプに分けている。Multi とは、複数質問、複数答弁の要約となっている場合 ([1] や [2] が含まれる要約の場合) のことである。Single とは、要約結果に [1] や [2] が含まれない要約である。

```
{
  "ID": "Segmentation-2018-JA-FormalTraining-00020",
  "Prefecture": "東京都",
  "Date": "24-2-28",
  "Meeting": "平成24年第1回定例会",
  "MainTopic": "がれき処理を迅速にし次計画を<br>外国人看護師候補者等に支援を",
  "SubTopic": "盲ろう者支援",
  "QuestionSpeaker": "中嶋義雄 (公明党)",
  "QuestionSummary": "[1] 区市町村と連携を。[2] 国にナショナルセンター設置を強く求めるべき。",
  "AnswerSpeaker": "福祉保健局長",
  "AnswerSummary": "[1] 研修会等で支援センター活用を働きかける。[2] 支援策充実を働きかける。",
  "QuestionStartingLine": 23037,
  "QuestionEndingLine": 23048,
  "AnswerStartingLine": 23234,
  "AnswerEndingLine": 23239
}
```

図 2: Segmentation Task の JSON 形式データの例

表 3: Summarization Task のデータ形式

表記	説明
Prefecture	都道府県名
Date	日付
Meeting	会議名 (議会だよりの記載)
Speaker	発言者
StartingLine	開始行
EndingLine	終了行
Main topic	メインピック
Sub topic	サブピック
Summary	要約文
Length	文字数
Source	原文

表 4: Summarization のデータ数

Training data	Test data	
	Single	Multi
596	79	67

²<http://www.gikai.metro.tokyo.jp/newsletter/>

```

{
  "ID": "Summarization-2018-JA-FormalTraining-00586",
  "Prefecture": "東京都",
  "Date": "24-11-30",
  "Meeting": "平成24年第4回定例会",
  "Speaker": "福祉保健局長",
  "MainTopic": "有効な子宮頸がん検診の検討を<br>若年性認知症の支援体制構築を",
  "SubTopic": "子宮頸がん",
  "StartingLine": 43700,
  "EndingLine": 43701,
  "Summary": " 国の指針で2年に1回細胞診を実施。HPV検査の実施は国の動向を注視。",
  "Length": "50字以内",
  "Source": "次に、子宮頸がんの有効な検診のあり方についてでございますが、国は、自治体を実施するがん検診について、死亡率減少効果が科学的に証明された検診方法を指針で示しており、子宮がん検診は、二十歳以上を対象に、二年に一回、細胞診を実施することとされております。\\n現在、国は、子宮がん検診として、従来の細胞診に加え、ヒトパピローマウイルスの検査を実施することについて検討しており、都としても国の動向を注視してまいります。\\n"
}

```

図 3: Summarization Task の JSON 形式データの例

2.3 Classification Task

Classification Task では、課題文と会議録に含まれる発言を比較し、関連、事実検証可能性、立場について分類することで、根拠のある意見をみつけることを目的とする。

課題文とは、政治課題と意見が含まれており「政治家の発言³」と比較する文のことである。図 4 に示した課題文「築地市場を豊洲に移転するべきである」と分類例とともに説明する。関連 (Relevance) とは「政治家の発言」と「課題文」との間にかかわりがあることである。事実検証可能 (Fact-checkability) とは「政治家の発言」に施設名や日付や金額などの具体的な表現が含まれ、Fact Check(事実の検証) ができることである。立場 (Stance) とは「課題文」に対する発言者の立ち位置である。Support は課題文の内容を支持していること、Against は課題文の内容を支持していないこと、Other は支持・不支持がはっきりしない場合や中立の場合である。

本タスクの入力、出力、評価は、次の通りである。

入力: 政策が記述された「課題文」と議会議録中の「発言文 (一文のみ)」
 出力: 3つのサブクラス「関連 (有り, 無し)」
 「事実検証 (可能, 不可能)」
 「立場 (支持, 不支持, どちらでもない)」を分類する。
 評価: 分類クラスの一致率

表 5 にデータ形式を示す。図 6 に Formal run の JSON 形式の例を示す。

データ作成は、議会で議論されている課題を選び、課題文のキーワードを含む発言を 47 都道府県議会か

³発言は一文のみとしている。

ら抽出した後に、クラウドソーシングを利用して、3人以上にラベル付けを依頼した。表 6 に課題文と注釈者と Training data 数と Test data 数を示す。

表 5: Classification Task のデータ形式

表記	説明
ID	識別子
Topic	政策、
Utterance	発言文
Relevance	1 … 関連あり 0 … 関連なし
Fact-checkability	1 … 事実検証可能 0 … 事実検証不可能
Stance	1 … Support (支持) 2 … Against (不支持) 0 … Other (どちらでもない)
Class	1 … Fact-checkable Support 2 … Fact-checkable Against 0 … Other

表 6 に Classification task の 14 の課題文と注釈者数を示す。注釈者の総数は 46 人であり、異なり注釈者数は 20 人である。

3 おわりに

本稿では、Formal run の 3つのタスク (Segmentation, Summarization, Classification) の意義と Formal Run Dataset の構築方法について述べた。

謝辞

本研究は JSPS 科研費 JP16H02912 および平成 30 年度国立情報学研究所公募型共同研究の助成を受けています。本タスクの設計等にご助言いただいた理化学

文ID	発言文（文脈のない一文） ※ 連続した発言ではありません。	関連	事実検証可能	立場
1	豊洲は、新市場移転により千客万来施設ができるなど、今後、観光客の集客が大いに期待できるエリアであります。	有	可能	支持
2	豊洲の新市場予定地では108箇所で液状化が発生し、築地市場の移転先としてふさわしくないことが重ねて証明されました。	有	可能	不支持
3	新銀行東京や築地市場の移転問題は非の立場です。	有	不可能	不支持
4	このような中、東京都はこの八月三十日に、豊洲の土壌汚染対策工事として、ゼネコン系の三つのJVと合計約五百四十二億円の契約を交わしています。	有	可能	どちらでもない
5	豊洲新市場への移転を希望する事業者の不安を払拭すべく、移転資金や運転資金、移転後の新たな事業展開に必要な資金の手当てなど、経営支援策を講じてまいります。	有	不可能	どちらでもない
6	お正月五日の築地の初競りで史上最高となる一億五千五百四十万で落札された大間マグロはブランド化成功の代表例であります	無	可能	どちらでもない
7	京築地域ではどのように三次医療が確保されているのかお聞かせ願います	無	不可能	どちらでもない

図 4: 「築地市場を豊洲に移転するべきである」に対する関連, 事実検証, 立場の例

表 6: Classification の課題文とデータ数

番号	課題文	注釈者数	Training	Test
1	カジノを含む統合型リゾートを推進するべきである	3	1,165	365
2	集団的自衛を認めるべきである	3	1,515	496
3	ハッ場ダムを建設を進めるべきである	3	886	292
4	高齢者への医療助成を増やすべきである	3	664	222
5	私学助成を推進するべきである	3	502	169
6	中京圏構想を推進するべきである	5	226	48
7	オスプレイを配備する	3	1,174	396
8	特定秘密保護法案を進めるべきである	3	462	163
9	道州制を導入するべきである	3	1,152	388
10	子ども医療費を無料化するべきである	5	241	73
11	正規の教員を増やすべきである	3	345	119
12	生活保護の基準額を引き下げるべきである	3	827	294
13	東京にオリンピックを招致するべきである	3	834	281
14	行政の判断で空き家を処理できるようにするべきである	3	298	106

関連	事実検証	立場	Class
有	可能	支持	Fact-checkable Support
有	可能	不支持	Fact-checkable Against
有	可能	どちらでもない	Other
有	不可能	支持	Other
有	不可能	不支持	Other
有	不可能	どちらでもない	Other
無	可能	支持	Other
無	可能	不支持	Other
無	可能	どちらでもない	Other
無	不可能	支持	Other
無	不可能	不支持	Other
無	不可能	どちらでもない	Other

図 5: 関連, 事実検証, 立場と class の関係

```
{
  "ID": "Classification-2018-JA-FormalSample-00001",
  "Topic": "築地市場を豊洲に移転するべきである",
  "Utterance": "豊洲は、新市場移転により千客万来施設ができるなど、今後、観光客の集客が大いに期待できるエリアであります。",
  "Relevance": 1,
  "Fact-checkability": 1,
  "Stance": 1,
  "Class": 1
}
```

図 6: Classification Task の JSON 形式データの例

参考文献

[1] Hideyuki Shibuki, Kotaro Sakamoto, Yoshinobu Kano, Teruko Mitamura, Madoka Ishioroshi, Kelly Y. Itakura, Di Wang, Tatsunori Mori, and Noriko Kando. Overview of the ntcir-11 qa-lab task. Proceedings of the 11th NTCIR Conference, 2014.

研究所関根聡氏, 東北大学乾健太郎教授, 横浜国立大学森辰則教授に感謝いたします。会議録の著作権等にご助言いただいた情報セキュリティ大学院大学湯浅壺道教授に感謝いたします。