

NTCIR 14

QA Lab PolilInfo  
第5回説明会

1208会議室@NII

2019年04月23日15:00-17:30

# NTCIR-14 QA Lab PoliInfo

<https://poliinfo.github.io/>

## Organizers:

木村泰知 (小樽商科大学, 理化学研究所AIP)

渋谷英潔 (横浜国立大学)

高丸圭一 (宇都宮共和大学)

内田ゆず (北海学園大学)

乙武北斗 (福岡大学)

石下円香 (国立情報学研究所)

阪本浩太郎 (横浜国立大学, 国立情報学研究所)

三田村照子 (カーネギーメロン大学)

神門典子 (国立情報学研究所, 総合研究大学院大学)

## Advisers:

森辰則 (横浜国立大学)

湯浅壘道 (情報セキュリティ大学院大学)

関根聡 (理化学研究所AIP)

乾健太郎 (東北大学, 理化学研究所AIP)

# 第5回説明会の概要

- Formal Run の結果について
  - QALab-PoliInfoのデータ分析
    - ▶ タイトル「注釈者のクラスタリングに基づく議会発言の事実検証可能性の推定」
    - ▶ 発表者 白土大樹、秋葉友良（豊橋技術科学大学）
  - The 14th NTCIR Conference について
  - NTCIR15 QALab-PoliInfoに向けて
  - 意見交換
- 
- 懇親会 (18:30)

Formal Run の結果について

# Task description

## 1. Segmentation Task

- ▶ 入力：質問と答弁の「発言の引用の対(つい)」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」）

## 2. Summarization Task

- ▶ 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」

## 3. Classification Task

- ▶ 入力：政策が記述された「課題文」と議会会議録中の「発言文（一文のみ）」
- ▶ 出力：3値に分類「事実検証可能な根拠を伴う賛成」「事実検証可能な根拠を伴う反対」「その他」

### ✓ 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)」

# 参加チームと提出数

Table 5. Number of submitted runs

Team ID	Dry run			Formal run		
	Segmentation	Summarization	Classification	Segmentation	Summarization	Classification
FU01	-	-	1	-	-	3
FU02	-	-	1	-	-	2
KitAi	-	-	-	-	2	-
TTECH	-	1	4	-	1	10
nami	11	-	-	11	-	-
nagoy	-	1	-	-	1	-
akbl	1	2	1	3	2	1
ibrk	-	-	1	-	-	2
RICT	1	-	1	5	-	7
STARS	-	-	4	-	-	4
tmcit	-	-	1	-	-	6
KSU	2	1	-	4	6	8
CUTKB	-	-	-	-	-	1
LisLb	-	-	-	-	1	1
TO	1	1	-	1	1	-
Sum	16	6	14	24	14	45

# Segmentation

# Segmentation task 開始行について

## ● 開始行の決定方法

- ▶ まず、開始行を説明する前に「都議会会議録」と「都議会だより」の関係について説明します

東京都議会 Tokyo Metropolitan Assembly

サイトマップ モバイル English

文字サイズ 拡大 標準 キーワードを入力

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 > 297号 > 代表質問(吉田信夫)

防災、放射能対策を強化せよ  
暮らしと福祉を守る手立てを



吉田信夫（日本共産党）

防災

（1）東京の津波対策の遅れをどう認識し、対応するのか。（2）耐震化されていない河川堤防は、東部地域で68キロ残る。強化策を。（3）超高層マンションの長周期地震動対応に支援を。（4）また、室内の安全対策を図るべき。

**知事** （1）新たな被害想定の見直しを進めており、津波対策に万全を期していく。

**東京都技監** （2）技術検証委員会等で耐震性向上策を検討。これを踏まえ新整備計画を策定。

**都市整備局長** （3）対策を国に要望した。

**消防総監** （4）専門委員会を設置し審議した。今後も、総合的な室内安全対策を推進する。



# Segmentation task 開始行について

## ● 開始行の決定方法

### ▶ 「東京都議会だより」と「都議会会議録」の関係について

東京都議会 Tokyo Metropolitan Assembly

サイトマップ モバイル English

文字サイズ 拡大 標準 キーワードを入力

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 > 297号 > 代表質問(吉田信夫)

Main topic

防災、放射能対策を強化せよ  
暮らしと福祉を守る手立てを

質問者

吉田信夫（日本共産党）

Sub topic

質問要約

答弁者

防災

(1) 東京の津波対策の遅れをどう認識し、対応するのか。(2) 耐震化されていない河川堤防は、東部地域で68キロ残る。強化策を。(3) 超高層マンションの長周期地震動対応に支援を。(4) また、室内の安全対策を図るべき。

知事 (1) 新たな被害想定の見直しを進めており、津波対策に万全を期していく。

東京都技監 (2) 技術検証委員会等で耐震性向上策を検討。これを踏まえ新整備計画を策定。

都市整備局長 (3) 対策を国に要望した。

消防総監 (4) 専門委員会を設置し審議した。今後も、総合的な室内安全対策を推進する。

答弁要約

# Segmentation のJsonフォーマット(Format run バージョン)

## ● Segmentation Task

- ▶ **質問と答弁の要約**を与えます
- ▶ 質問の範囲と答弁の範囲を答えてもらいます

- ✓ 質問の開始行
- ✓ 質問の終了行
- ✓ 答弁の開始行
- ✓ 答弁の終了行

表記	説明
ID	識別子 タスク名-開催年-言語-タスク種別-連番
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

# Segmentation のJsonフォーマット(Format run バージョン)

```
{
  "ID":"Segmentation-2018-JA-FormalSample-00001",
  "Prefecture":"東京都",
  "Date":"23-6-23",
  "Meeting":"平成23年_第2回定例会",
  "MainTopic":"東京の防災力向上は待ったなし<br>被災地復興には総合的
  取組を",
  "SubTopic":"都政運営",
  "QuestionSpeaker":"村上英子（自民党）",
  "QuestionSummary":"(1) 国難のもとにおける東京の役割について所見
  を。(2) 「10年後の東京」計画改定の視点は。(3) 都民への責任果た
  し得る財政運営を。(4) 被災地復興支援と首都東京の防災力向上に今後
  どう取り組むのか、所見を。","AnswerSpeaker":"知事本局長",
  "AnswerSummary":"(2) 高度防災機能備えた都市へ取組強化等で安全・安心社会をつくり、節電意識徹底等で環境と経済が両立した都市実現
  等。",
  "QuestionStartingLine":763,
  "QuestionEndingLine":793,
  "AnswerStartingLine":1240,
  "AnswerEndingLine":1243
},
```

埋める

表記	説明
ID	識別子 タスク名ー開催年ー言語ータスク種別ー連番
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

# Segmentation task のデータセット

Training

質問	298
答弁	298

298セット

質問と答弁を対(つい)としているので  
298セットとなる

Test

質問	83
答弁	83

83セット

# Segmentationの結果

Table 10. result of Segmentation task in Formal run

	<i>R</i>	<i>P</i>	<i>F</i>
nami-01	0.814 (1,433/1,761)	<u>0.940</u> (1,433/1525)	0.872
nami-02	0.864 (1,521/1,761)	0.851 (1,521/1,788)	0.857
nami-03	0.984 (1,733/1,761)	0.499 (1,733/3,475)	0.662
nami-04	0.639 (1,125/1,761)	0.805 (1,125/1,398)	0.712
nami-05	0.553 (973/1,761)	0.931 (973/1,045)	0.694
nami-06	0.655 (1,153/1,761)	0.657 (1,153/1,754)	0.656
nami-07	0.797 (1,404/1,761)	0.933 (1,404/1,505)	0.860
nami-08	0.831 (1,464/1,761)	0.932 (1,464/1,570)	0.879
nami-09	0.875 (1,541/1,761)	0.843 (1,541/1,827)	0.859
nami-10	0.993 (1,749/1,761)	0.464 (1,749/3,769)	0.632
nami-11	<u>1.000</u> (1,761/1,761)	0.112 (1,761/15,765)	0.201
akbl-01	0.768 (1,352/1,761)	0.538 (1,352/2,515)	0.633
akbl-02	0.847 (1,492/1,761)	0.455 (1,492/3,282)	0.592
akbl-03	0.656 (1,155/1,761)	0.519 (1,155/2,227)	0.580
RICT-01	0.882 (1,554/1,761)	0.909 (1,554/1,709)	<u>0.895</u>
RICT-02	0.856 (1,507/1,761)	0.889 (1,507/1,695)	<u>0.872</u>
RICT-03	0.853 (1,503/1,761)	0.780 (1,503/1,926)	0.815
RICT-04	0.780 (1,374/1,761)	0.746 (1,374/1,842)	0.763
RICT-05	0.936 (1,648/1,761)	0.712 (1,648/2,314)	0.809
KSU-01	0.779 (1,372/1,761)	0.243 (1,372/5,643)	0.370
KSU-02	0.759 (1,337/1,761)	0.268 (1,337/4,998)	0.396
KSU-03	0.820 (1,444/1,761)	0.661 (1,444/2,185)	0.732
KSU-04	0.797 (1,403/1,761)	0.922 (1,403/1,521)	0.855
TO-01	0.354 (623/1,761)	0.898 (623/694)	0.508

# Summarization

# SummarizationのJsonフォーマット(Format run バージョン)

## ● Summarization Task

- ▶ 都議会会議録データを利用します
  - ✓ Pref13\_tokyo.json
- ▶ トピックについて
  - ✓ 「メインピック」「サブトピック」の追加
- ▶ 会議録データの対象範囲を明記します
  - ✓ 「開始行」「終了行」の追加

表記	説明
ID	識別子 <small>タスク名-開催年-言語-タスク種別-連番</small>
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	サブトピック
Summary	要約文
Length	文字数
Source	原文

# SummarizationのJsonフォーマット(Format run バージョン)

```
{  
  "ID": "Summarization-2018-JA-Dry-00002",  
  "Prefecture": "東京都",  
  "Date": "23-6-23",  
  "Meeting": "平成23年第2回定例会",  
  "Speaker": "石原慎太郎",  
  "StartingLine": 541,  
  "EndingLine": 550,  
  "Main topic": "東京の総合防災力を更に高めよ<br>環境に配慮した都市づくりを",  
  "Sub topic": "原子力発電所",  
  "Summary": "  原子力利用に当たって安全の確保は当然。信頼性失ってきた姿勢を反省し再考すべき。  
  同時にエネルギー確保は国家・社会の存立に直結、戦略を速やかに構え直す必要あり。",  
  "Length": "100字以内",
```

埋める

```
"Source": "次いで、原子力発電の安全性についてであります。今回の福島原発の事故は、ただいまの  
質問にあった四年前、柏崎原発が被災した際、東京都が国に安全性確保を要求した云々の話とは全く次元の  
違う問題であります。¥¥¥先日、海江田経済産業大臣は、原発の短期的な安全については確認されたと表明  
しております。¥¥¥みずからの民主党政権が原発を管轄し、検査もしているのですから、安全性については  
国家に聞いたらいかがですか。¥¥¥改めて申し上げることもなく、原子力の利用に当たっても、安全の確保  
は当然であります。¥¥¥なぜフランスでできている安全管理が日本にできないのか。¥¥¥インシデントすら  
隠して信頼性を失ってきたこれまでの姿勢を反省し、再考すべきであると思います。¥¥¥同時に、エネル  
ギー確保は、国家、社会の存立に直結するがゆえに、エネルギー戦略も速やかに構え直す必要があります。  
¥¥¥豊富な電力が安定的に供給されるからこそ、我が国を支えるものづくりがここまで発展できたわけであ  
りまして、新しいITの技術が急速に進歩してまいりました。¥¥¥かつて、完全な管理技術を前提とすれば、  
東京湾に立派な原子力発電をつくってもいいと発言したことがございますが、当然のことです。¥¥¥  
日本の頭脳、心臓部の将来を、一段の発展を考えたとき、その裏づけとなる電力をほかに依存し続けること  
は不合理だと考えます。¥¥¥"
```

```
},
```

表記	説明
ID	識別子 タスク名ー開催年ー言語ータスク種別ー連番
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	サブトピック
Summary	要約文
Length	文字数
Source	原文



# Summarization task のデータセット

Training

質問	298
答弁	298

} 596セット

Test

Single	79
Multi	67

} 146セット

# Summarizationの結果(1/2)

「○」を2点、「△」を1点、「×」を0点として計算し、評価の観点ごとに平均化

表 1: 人手評価のスコア (スコアの最大値は 2)

	all-topic				single-topic				multi-topic			
	内容		表現	全体	内容		表現	全体	内容		表現	全体
	▲=0	▲=2			▲=0	▲=2			▲=0	▲=2		
akbl-1	0.722	1.005	1.833	0.826	0.708	1.009	1.844	0.849	0.739	1.000	1.821	0.799
akbl-2*	0.707	1.000	1.837	0.793	—	—	—	—	0.707	1.000	1.837	0.793
KitAi-1	0.856	1.134	1.732	0.912	0.953	1.170	1.660	0.995	0.745	1.092	1.815	0.815
KitAi-2	0.788	1.035	1.308	0.667	0.849	1.028	1.340	0.722	0.717	1.043	1.272	0.603
KSU-1	0.043	0.043	1.955	0.048	0.052	0.052	1.934	0.057	0.033	0.033	1.978	0.038
KSU-2	0.076	0.121	1.745	0.071	0.080	0.156	1.722	0.104	0.071	0.082	1.772	0.033
KSU-3	0.091	0.157	1.715	0.104	0.104	0.179	1.731	0.156	0.076	0.130	1.696	0.043
KSU-4	0.111	0.167	1.419	0.093	0.118	0.193	1.420	0.132	0.103	0.136	1.418	0.049
KSU-5	0.048	0.078	1.692	0.048	0.057	0.085	1.726	0.057	0.038	0.071	1.652	0.038
KSU-6	0.078	0.169	1.535	0.091	0.085	0.151	1.542	0.094	0.071	0.190	1.527	0.087
LisLb-1	0.720	0.942	1.237	0.591	0.722	0.920	1.349	0.684	0.717	0.967	1.109	0.484
nagoy-1	0.886	1.104	1.619	0.899	0.953	1.179	1.642	1.028	0.810	1.016	1.592	0.750
TO-1**	0.504	0.846	1.763	0.551	0.464	0.794	1.778	0.521	0.550	0.905	1.746	0.586
TTECH-1	0.290	0.644	1.783	0.402	0.274	0.575	1.755	0.401	0.310	0.723	1.815	0.402
平均	0.423	0.603	1.655	0.435	0.387	0.535	1.532	0.414	0.406	0.599	1.646	0.394

\*akbl-2 は single-type 未提出

\*\*TO-1 はタスクオーガナイザによる結果

# Summarizationの結果(2/2)

表 2: ROUGE のスコア (all-topic を対象)

		recall							F-measure						
		N1	N2	N3	N4	L	W1.2	SU4	N1	N2	N3	N4	L	W1.2	SU4
表 層 形	akbl-1	.400	.173	.113	.076	.345	.189	.157	.361	.156	.102	.068	.310	.167	.185
	akbl-2	.326	.124	.080	.057	.269	.147	.112	.320	.119	.077	.055	.262	.141	.144
	KitAi-1	.440	.185	.121	.085	.375	.217	.179	.357	.147	.096	.067	.299	.168	.188
	KitAi-2	.390	.174	.113	.078	.320	.200	.154	.343	.154	.101	.069	.281	.173	.176
	KSU-1	.158	.028	.009	.002	.147	.043	.071	.210	.039	.013	.004	.196	.059	.107
	KSU-2	.185	.043	.021	.014	.167	.063	.080	.230	.056	.027	.017	.209	.080	.116
	KSU-3	.172	.036	.008	.002	.157	.050	.075	.211	.043	.011	.003	.192	.062	.106
	KSU-4	.171	.044	.013	.002	.153	.055	.072	.219	.056	.017	.003	.195	.072	.106
	KSU-5	.227	.029	.010	.002	.195	.064	.089	.231	.029	.010	.003	.196	.065	.110
	KSU-6	.221	.038	.013	.004	.187	.065	.086	.230	.038	.012	.004	.192	.067	.108
	LisLb-1	.251	.120	.079	.058	.211	.132	.103	.226	.107	.071	.051	.188	.115	.118
	nagoy-1	.459	.200	.131	.089	.394	.229	.186	.361	.151	.097	.064	.305	.169	.192
	TO-1	.267	.093	.061	.045	.230	.117	.105	.272	.086	.052	.036	.233	.110	.133
	TTECH-1	.278	.060	.035	.020	.216	.092	.096	.240	.055	.031	.018	.187	.079	.111
原 形	akbl-1	.415	.184	.122	.083	.357	.203	.164	.375	.165	.110	.074	.322	.179	.195
	akbl-2	.339	.135	.089	.064	.279	.158	.119	.333	.129	.085	.063	.272	.152	.153
	KitAi-1	.458	.199	.134	.096	.389	.234	.188	.373	.159	.106	.075	.311	.182	.199
	KitAi-2	.399	.179	.118	.082	.326	.208	.158	.351	.160	.106	.074	.286	.180	.181
	KSU-1	.161	.028	.010	.002	.148	.044	.071	.214	.040	.013	.004	.197	.061	.108
	KSU-2	.187	.044	.021	.014	.170	.064	.081	.233	.057	.027	.017	.212	.082	.117
	KSU-3	.175	.036	.008	.002	.159	.052	.075	.217	.044	.011	.003	.196	.065	.108
	KSU-4	.174	.045	.014	.002	.155	.056	.073	.222	.058	.018	.003	.197	.073	.107
	KSU-5	.230	.029	.010	.002	.199	.066	.090	.236	.030	.010	.003	.201	.067	.112
	KSU-6	.226	.040	.013	.004	.189	.066	.087	.235	.039	.012	.004	.195	.069	.109
	LisLb-1	.261	.125	.084	.061	.218	.139	.106	.235	.112	.075	.055	.195	.121	.122
	nagoy-1	.479	.217	.145	.101	.412	.247	.197	.377	.165	.108	.074	.319	.184	.205
	TO-1	.273	.097	.065	.048	.233	.121	.107	.277	.089	.056	.039	.236	.114	.136
	TTECH-1	.289	.064	.037	.022	.222	.097	.099	.251	.058	.033	.019	.193	.084	.114
内 容 語	akbl-1	.256	.113	.065	.034	.247	.124	.148	.224	.098	.056	.031	.216	.100	.158
	akbl-2	.200	.094	.051	.032	.189	.095	.109	.188	.089	.049	.031	.178	.087	.127
	KitAi-1	.285	.145	.090	.050	.278	.154	.180	.224	.115	.071	.042	.217	.107	.170
	KitAi-2	.254	.126	.083	.053	.247	.131	.156	.214	.109	.069	.046	.208	.106	.159
	KSU-1	.048	.001	.000	.000	.047	.007	.032	.059	.001	.000	.000	.058	.009	.043
	KSU-2	.069	.014	.000	.000	.067	.019	.043	.083	.015	.000	.000	.081	.022	.059
	KSU-3	.041	.002	.000	.000	.041	.007	.027	.050	.002	.000	.000	.050	.008	.036
	KSU-4	.050	.002	.000	.000	.048	.008	.031	.064	.003	.000	.000	.061	.011	.044
	KSU-5	.067	.002	.000	.000	.062	.013	.041	.063	.003	.000	.000	.057	.011	.043
	KSU-6	.053	.003	.000	.000	.051	.008	.034	.051	.003	.000	.000	.049	.009	.037
	LisLb-1	.171	.083	.044	.026	.160	.088	.106	.140	.068	.036	.023	.130	.065	.102
	nagoy-1	.326	.164	.094	.046	.315	.168	.201	.249	.123	.067	.036	.239	.110	.187
	TO-1	.116	.055	.035	.012	.111	.056	.070	.106	.042	.023	.011	.101	.042	.076
	TTECH-1	.088	.028	.015	.007	.082	.033	.050	.076	.024	.012	.006	.071	.027	.054

ROUGE-N1、-N2、-N3、-N4、-L、-W1.2、-SU4 の7 通りの手法を用い、各手法に対し、再現率を用いた場合とF 値を用いた場合のスコアを計算した。

# Classification

# ラベル付けについて

## ● 目的

▶ 政治家の発言から「事実検証可能な根拠を伴う意見」を見つけることです。

## ● 作業

▶ 「課題文」と「政治家の発言(一文のみ)」を読み、次の3つのラベルを付けてもらいます。

1. 「関連」 ... 有り or 無し
2. 「事実検証可能」 ... 可能 or 不可能
3. 「立場」 ... Support , Against or Other

# ラベル付けについて

## ▶ 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(一文のみ)」と比較する文のことです。

- 例「**築地市場を豊洲に移転すべきである**」

## ▶ 関連(Relevance)とは

✓ 「政治家の発言(一文のみ)」と「課題文」との間にかかわりがあることです。

- 関連のある例 「新銀行東京や**築地市場の移転**問題は非の立場です。」
- 関連のない例 「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」

## ▶ 事実検証可能(Fact-checkability)とは

✓ 「政治家の発言(一文のみ)」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)ができることです。

- 事実検証できる例 「このような中、東京都はこの**八月三十日**に、**豊洲の土壤汚染対策工事**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。」
- 事実検証できない例 「私は、**築地の場外市場**のようなあの温かみのある、あの何ともいえないにぎわいをつくるのが、都の責任の一つであると考えます」

## ▶ 立場(Stance)とは

✓ 「課題文」に対する発言者の立ち位置です。Support(支持), Against(不支持), Neutral(中立,どちらでもない) から一つ選びます。

- Support は、課題文の内容を支持していること
  - Support の例 「**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。」
- Against は、課題文の内容を支持していないこと
  - Against の例 「**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。**」
- Otherは、支持、不支持がはっきりしない場合や中立の場合
  - (発言者の言葉ではない) 引用や報告は意見ではありません、

# 関連、事実検証可能、立場のラベルをつける例 (1/4)

課題文「築地市場を豊洲に移転するべきである」と下記の発言文を読み、3つのラベルをつける

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。			
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。			
3	新銀行東京や築地市場の移転問題は非の立場です。			
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。			
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。			
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			
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			

# 関連、事実検証可能、立場のラベルをつける例 (2/4)

課題文「築地市場を豊洲に移転するべきである」と下記の発言文を読み、3つのラベルをつける

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、 <u>新市場移転</u> 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		
2	豊洲の新市場予定地では108箇所で液状化が発生し、 <u>築地市場の移転先</u> としてふさわしくないことが重ねて証明されました。	有		
3	新銀行東京や <u>築地市場の移転問題</u> は非の立場です。	有		
4	このような中、東京都はこの八月三十日に、 <u>豊洲の土壌汚染対策工事</u> 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		
5	<u>豊洲新市場への移転</u> 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		
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		
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		



# 関連、事実検証可能、立場のラベルをつける例 (3/4)

課題文「築地市場を豊洲に移転するべきである」と下記の発言文を読み、3つのラベルをつける

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	
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	
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	
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	
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	
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	
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	

# 関連、事実検証可能、立場のラベルをつける例 (4/4)

課題文「築地市場を豊洲に移転するべきである」と下記の発言文を読み、3つのラベルをつける

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

# 関連、事実検証可能、立場 と Class の関係

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

# 課題文と文数と注釈者

14の課題文

Training data の異なり文数 約1万文

Training data の総文数 約3万文

注釈者総数 46人

異なり注釈者数 20人

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	八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	中京都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にするべきである	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

# Classificationの評価

## 略語の意味

- 正解のみなし方
  - [アルファベット1文字] : アルファベットは作成者IDを表す. その作成者による成果情報を用いたGold Standard.
  - N1 : 作成者が1人でも付与したラベルを正解とみなす.
  - N2 : 作成者が2人以上付与したラベルを正解とみなす.
  - N3 : 作成者が3人以上付与したラベルを正解とみなす. 3人以上付与したラベルがない場合は正解なしとして, 評価対象から除外する.
  - SC : 作成者が何人付与したかを正解スコアの重みとして用いる, (例えば, 0が2名, 1が1名の場合, 出力0に対しては2点, 出力1に対しては1点という評価をする)
- 評価対象
  - R1 : Relevance
  - FC : Fact-checkability
  - St : Stance
  - Cl : Class

# Classificationの結果

	A	support			against			other		
		R	P	F	R	P	F	R	P	F
FU01-01	0.624	0.417	0.057	0.100	0.076	0.041	0.053	0.648	0.938	0.766
FU01-02	0.373	0.731	0.057	0.106	0.183	0.045	0.072	0.362	0.943	0.523
FU01-03	0.909	0.089	0.164	0.115	0.008	0.020	0.011	0.970	0.936	0.953
FU02-01	0.842	0.027	0.040	0.032	0.095	0.033	0.049	0.899	0.933	0.916
FU02-02	0.840	0.073	0.063	0.068	0.069	0.030	0.042	0.895	0.933	0.914
TTECH-01	0.923	0.046	0.163	0.072	0.015	0.133	0.027	0.987	0.935	0.960
TTECH-02	0.896	0.260	0.252	0.256	0.221	0.199	0.209	0.943	0.947	0.945
TTECH-03	0.919	0.116	0.254	0.159	0.069	0.200	0.103	0.978	0.938	0.958
TTECH-04	0.921	0.043	0.134	0.065	0.015	0.133	0.027	0.985	0.934	0.959
TTECH-05	0.897	0.251	0.251	0.251	0.225	0.207	0.216	0.944	0.947	0.945
TTECH-06	0.918	0.132	0.269	0.177	0.080	0.206	0.115	0.976	0.939	0.957
TTECH-07	0.942	0.000	NaN	NaN	0.000	NaN	NaN	1.000	0.942	0.970
TTECH-08	0.942	0.000	NaN	NaN	0.000	NaN	NaN	1.000	0.942	0.970
TTECH-09	0.926	0.000	0.000	NaN	0.000	NaN	NaN	0.982	0.941	0.961
TTECH-10	0.942	0.000	NaN	NaN	0.000	NaN	NaN	1.000	0.942	0.970
akbl-01	0.923	0.118	0.344	0.176	0.034	0.097	0.050	0.983	0.939	0.960
ibrk-01	0.731	0.178	0.063	0.093	0.202	0.045	0.074	0.770	0.934	0.844
ibrk-02	0.731	0.178	0.063	0.093	0.202	0.045	0.074	0.770	0.934	0.844
RICT-01	0.933	0.000	NaN	NaN	0.000	NaN	NaN	1.000	0.933	0.965
RICT-02	0.932	0.002	0.091	0.004	0.004	0.111	0.008	0.998	0.933	0.964
RICT-03	0.893	0.118	0.145	0.130	0.111	0.117	0.114	0.949	0.940	0.944
RICT-04	0.894	0.114	0.143	0.127	0.111	0.117	0.114	0.950	0.939	0.944
RICT-05	0.933	0.000	NaN	NaN	0.000	0.000	NaN	1.000	0.933	0.965
RICT-06	0.933	0.000	NaN	NaN	0.000	NaN	NaN	1.000	0.933	0.965
RICT-07	0.932	0.084	0.440	0.141	0.042	0.407	0.076	0.994	0.937	0.965
STARS-01	0.933	0.000	NaN	NaN	0.000	NaN	NaN	1.000	0.933	0.965
STARS-02	0.889	0.002	0.002	0.002	0.000	NaN	NaN	0.953	0.933	0.943
STARS-03	0.889	0.002	0.002	0.002	0.000	NaN	NaN	0.953	0.933	0.943
STARS-04	0.889	0.002	0.002	0.002	0.000	NaN	NaN	0.953	0.933	0.943
tmcit-01	0.875	0.282	0.139	0.186	0.000	NaN	NaN	0.925	0.943	0.934
tmcit-02	0.893	0.239	0.160	0.192	0.000	NaN	NaN	0.946	0.942	0.944
tmcit-03	0.873	0.296	0.142	0.192	0.000	NaN	NaN	0.922	0.943	0.932
tmcit-04	0.879	0.319	0.161	0.214	0.000	NaN	NaN	0.928	0.944	0.936
tmcit-05	0.898	0.267	0.189	0.221	0.000	NaN	NaN	0.950	0.942	0.946
tmcit-06	0.878	0.292	0.148	0.196	0.000	NaN	NaN	0.927	0.943	0.935
KSU-01	0.932	0.075	0.579	0.133	0.008	0.056	0.014	0.995	0.937	0.965
KSU-02	0.932	0.071	0.689	0.129	0.008	0.042	0.013	0.995	0.937	0.965
KSU-03	0.934	0.071	0.738	0.130	0.008	0.083	0.015	0.998	0.937	0.967
KSU-04	0.934	0.071	0.738	0.130	0.008	0.083	0.015	0.998	0.937	0.967
KSU-05	0.932	0.075	0.579	0.133	0.019	0.111	0.032	0.995	0.937	0.965
KSU-06	0.932	0.071	0.689	0.129	0.019	0.088	0.031	0.995	0.937	0.965
KSU-07	0.934	0.071	0.738	0.130	0.011	0.100	0.020	0.997	0.937	0.966
KSU-08	0.934	0.071	0.738	0.130	0.011	0.100	0.020	0.997	0.937	0.966
CUTKB-04	0.025	0.000	NaN	NaN	1.000	0.025	0.049	0.000	NaN	NaN
LisLb-01	0.914	0.021	0.065	0.032	0.037	0.080	0.051	0.976	0.935	0.955

# Classificationの結果

	<i>A</i>	support			against			other		
		<i>R</i>	<i>P</i>	<i>F</i>	<i>R</i>	<i>P</i>	<i>F</i>	<i>R</i>	<i>P</i>	<i>F</i>
FU01-01	0.624	0.417	0.057	0.100	0.076	0.041	0.053	0.648	0.938	0.766
FU01-02	0.373	<u>0.731</u>	0.057	0.106	0.183	0.045	0.072	0.362	0.943	0.523
FU01-03	0.909	0.089	0.164	0.115	0.008	0.020	0.011	0.970	0.936	0.953
FU02-01	0.842	0.027	0.040	0.032	0.095	0.033	0.049	0.899	0.933	0.916
FU02-02	0.840	0.073	0.063	0.068	0.069	0.030	0.042	0.895	0.933	0.914
TTECH-01	0.923	0.046	0.163	0.072	0.015	0.133	0.027	0.987	0.935	0.960
TTECH-02	0.896	0.260	0.252	<u>0.256</u>	0.221	0.199	0.209	0.943	<u>0.947</u>	0.945
TTECH-03	0.919	0.116	0.254	0.159	0.069	0.200	0.103	0.978	0.938	0.958
TTECH-04	0.921	0.043	0.134	0.065	0.015	0.133	0.027	0.985	0.934	0.959
TTECH-05	0.897	0.251	0.251	0.251	0.225	<u>0.207</u>	<u>0.216</u>	0.944	<u>0.947</u>	0.945
TTECH-06	0.918	0.132	0.269	0.177	0.080	0.206	0.115	0.976	0.939	0.957
TTECH-07	<u>0.942</u>	0.000	NaN	NaN	0.000	NaN	NaN	<u>1.000</u>	0.942	<u>0.970</u>
TTECH-08	<u>0.942</u>	0.000	NaN	NaN	0.000	NaN	NaN	<u>1.000</u>	0.942	<u>0.970</u>
TTECH-09	0.926	0.000	0.000	NaN	0.000	NaN	NaN	0.982	0.941	0.961
TTECH-10	<u>0.942</u>	0.000	NaN	NaN	0.000	NaN	NaN	<u>1.000</u>	0.942	<u>0.970</u>
tbl-01	0.923	0.118	0.244	0.176	0.034	0.007	0.050	0.983	0.930	0.960



# QALab-PoliInfoのデータ分析に関する報告

注釈者のクラスタリングに基づく  
議会発言の事実検証可能性の推定

白土大樹、秋葉友良  
(豊橋技術科学大学)



The 14th NTCIR Conference について

# The 14th NTCIR Conference について

## ● The 14th NTCIR Conference Program

▶ June 10 (Mon) – 13(Thu), 2019

▶ QA Lab-PoliInfo

✓ Overview(TO)

● 6/11 13:30-14:00

✓ Oral presentation (一部)

● 6/12 9:00-12:00

✓ Poster presentation (全員)

● 6/12 13:30-15:00

✓ Breakout session (全員)

● 6/12 15:00-16:30

● 今後について

NTCIR15 QALab-PoliInfoに向けて

# NTCIR15 QALab-PoliInfoに向けて

## ● タスクの修正 (案) について

▶ Classification task → Stance Classification

## ● 今後の取り組み (案) について

▶ SQuAD 2.0, HotpotQA, GLUEのような **Leaderboard** をつける

✓ 掲載条件：プログラムのGitHub公開？論文公開？

▶ データの拡張

✓ ディープラーニングを使えるデータ量にする。

## ● オーガナイザーの追加などの組織改編

▶ オーガナイザーの負担を軽減しつつ、参加者の成果が上がるようにする。

✓ 例えば「(全てではなく) 2つ以上のチームに所属して、作業してくれる方を募集する。」など…

- 企画・運営チーム
- データ基盤作成 チーム
- Segmentation チーム
- Summarization チーム
- Classification チーム
- 評価方法・ウェブ公開 チーム

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452
1	BERT + DAE + AoA (ensemble) Joint Laboratory of HIT and iFLYTEK Research	87.147	89.474
2	BERT + ConvLSTM + MTL + Verifier (ensemble) Layer 6 AI	86.730	89.286

# Important dates

Feb 20, 2018: QALab-PoliInfo Kickoff meeting in NII (room 1901, 1902)

Mar 20, 2018: NTCIR-14 Kickoff event in NII

Apr 19, 2018: 1st round table meeting in NII (room 1901, 1902)

May 31, 2018: 2nd round table meeting in NII (room 1208)

Jun 15, 2018: Dataset Release

Jul 30, 2018: Task Registration Due for Dry Run

Aug 6 - 9, 2018: Dry Run (Segmentation & Classification Tasks)

Aug 13 - 16, 2018: Dry Run (Summarization Task)

Aug 30, 2018: 3rd round table meeting in NII (room 2004 2005)

Oct 29, 2018: 4th round table meeting in NII

Nov 19, 2018: Task Registration Due for Formal Run (This is not required for Dry Run participants)

Nov 26 - 29, 2018: Formal Run (Segmentation & Classification Tasks)

Dec 3 - 6, 2018: Formal Run (Summarization Task)

Feb 1, 2019: Evaluation Result Release

Feb 1, 2019: Task overview paper release (draft)

Mar 15, 2019: Submission due of participant papers

May 1, 2019 Camera-ready participant paper due

Jun 10-13, 2019: NTCIR-14 Conference & EVIA 2019 in NII, Tokyo

# まとめ

- NTCIR14 QALab-PoliInfo Formal Run の 結果について
  - QALab-PoliInfoのデータ分析
    - ▶ タイトル 「注釈者のクラスタリングに基づく議会発言の事実検証可能性の推定」
    - ▶ 発表者 白土大樹、秋葉友良（豊橋技術科学大学）
  - The 14th NTCIR Conference について
  - NTCIR15 QALab-PoliInfoに向けて
  - 意見交換
- 
- 懇親会 (18:30)