

S1-08



人間とAI
が紡ぐ未来社会

シビックテックのための深層学習BERT を用いた社会課題タグ付け手法の開発

情報工学専攻 准教授 白松 俊

博士前期課程2年 長谷川 徳賢

概要

コラボレーション

自然言語処理×LODで社会課題解決を目指す人々の協働を支援

- **シビックテック (civic tech):** IT技術者と市民・行政の協働により社会課題にアプローチ
 - 技術者だけでは課題背景が捉えきれず、市民・行政だけでは技術の活用が難しい
- **Web記事への社会課題タグ付け:** 解決策を話し合う際、課題背景がわかる記事を探りたい
- **BERT (Bidirectional Encoder Representations from Transformer):**
 - 2018年秋に自然言語処理の精度を飛躍的に向上させた、深層学習による汎用言語モデル
- **LOD (Linked Open Data):** 世の中の様々な情報を構造化し、誰でも使えるように共有
 - **Wikidata:** Wikimedia財団による誰でも編集できるLOD (Wikipediaとも連動)
 - Wikidata上で社会問題を階層的に整理、タグとして利用 → 探索に適したタグ

特長

Wikidata上で階層化した社会課題タグ



階層構造を利用した課題と記事の探索を可能に

BERTによる自動タグ付けの精度向上

	正解率
先行研究 TF-IDF	0.81
パラグラフベクター	0.82
本研究 BERT	0.91

本技術が拓く心豊かな未来社会の姿

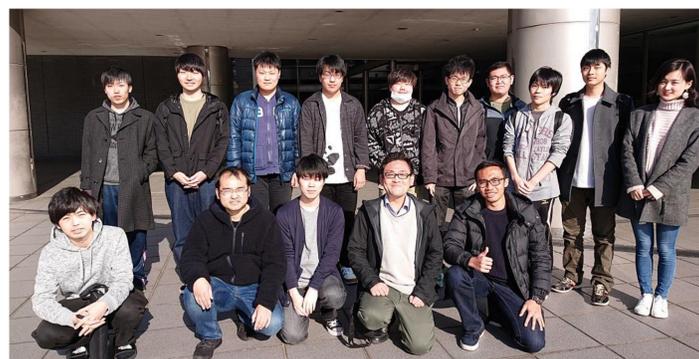
- **シビックテック活動での活用:** ハッカソン (参加型開発イベント) で活用
- **教育利用:** 学生が地域課題の解決策を考える授業や「調べ学習」で活用
- **公的な住民参加型議論での活用:** 他地域も含めた事例や社会的背景を参考にするため活用

今後の課題

- 深い階層のタグを自動タグ付けするための学習データをどう増やすか?
- 社会課題の解決を目指した取り組み (ミッション) も自動抽出したい
- 将来的には社会課題の因果関係もBERTで抽出し、解決策を考えるヒントに

求める連携先とメッセージ

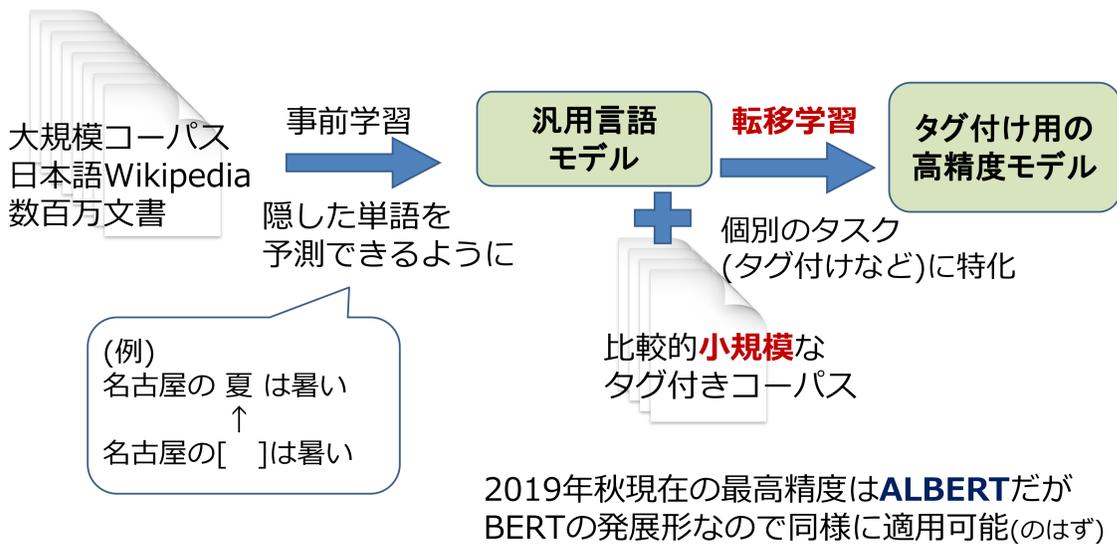
- CSR/CSV活動で貢献できる社会課題を探したい皆様
- シビックテックのような活動で社会貢献したい皆様
- 社会課題解決に事業として取り組んでいる皆様
- BERTとLODによるWebコンテンツ自動構造化技術で世の中の最新動向が見える化したい皆様



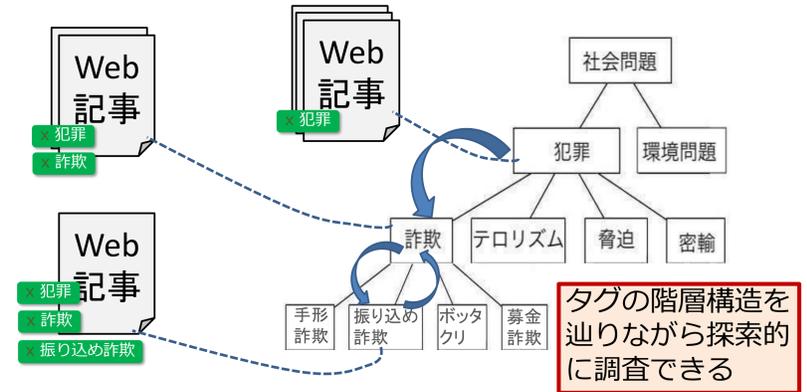
白松研究室

特長が発揮される仕組み

BERT: 様々な自然言語処理タスクで最高精度 (2018年秋)
(例: 固有表現抽出, 2文の含意関係判定, 質問応答判定, etc.)



階層的なタグを使った探索支援



- AIの知識源LODとして注目されている**Wikidata**
✓ 新たな社会課題のタグも市民参加により追加可能
- W3C標準 (**Web Annotation Data Model**) に準拠した自動タグ付け結果のLODを公開予定
✓ BERTの解析誤りも、市民参加により修正したい

技術の特長の根拠となる実験データ等

BERTを用いた社会課題タグ付け実験の精度 (第一/第二階層)

第一階層のタグ付け精度

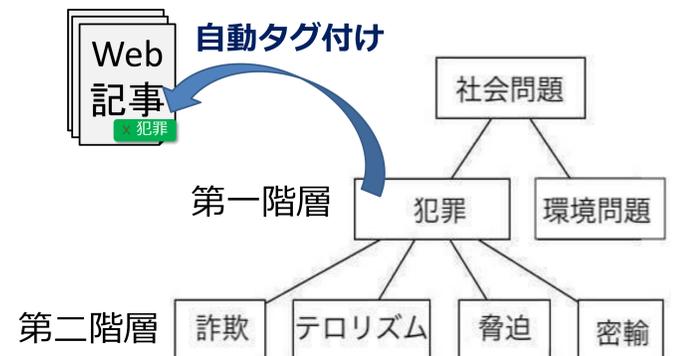
	precision	recall	f-score	support
災害	0.94	0.95	0.94	1773
犯罪	0.91	0.92	0.92	2571
⋮	⋮	⋮	⋮	⋮
戦争	0.92	0.84	0.88	667
accuracy			0.91	6269

“社会問題”に直接リンクのあるラベルのうち、下位クラス数が10以上かつ日本語Wikipediaの記事が存在するもの (10クラス分類)

第二階層のタグ付け精度

	precision	recall	f-score	support
テロリズム	0.90	0.90	0.90	42
拷問	0.91	0.85	0.88	60
⋮	⋮	⋮	⋮	⋮
詐欺	0.99	0.93	0.96	161
accuracy			0.91	785

第一階層と同様に“犯罪”に直接リンクのあるラベルのうち、下位クラス数が5以上かつ日本語Wikipediaの記事が存在するもの (7クラス分類)



Wikidataにおける階層構造の例

試作品の状況

提示可

※提供の際は諸手続が必要となるため、下記問合せ先までご連絡願います。

研究フェーズ



文献・特許の情報

- Masaru Watanabe, Shun Shiramatsu, Yasuaki Goto. "Tag-based Approaches to Sharing Background Information regarding Social Problems towards Facilitating Public Collaboration," in *eGose '17 Proceedings of the International Conference on Electronic Governance and Open Society: Challenges in Eurasia*, pp. 113-118, St. Petersburg, Russia, 2017.
- 成瀬雅人, 白松俊, 松島格也. "議会議事録の自動タグ付けに基づく探索的閲覧システムの試作とそのユーザビリティ評価", 第80回情報処理学会全国大会, 2ZB-07, 早稲田大学西早稲田キャンパス, 2018.
- 白松俊, Teemu Tossavainen, 大園忠親, 新谷虎松. "社会課題とその解決目標のLinked Open Data化による目標マッチングサービスの開発.", 人工知能学会論文誌, 31(1), pp. LOD-C_1-11, 2016.

【お問合せ】 名古屋工業大学 産学官金連携機構

〒466-8555 名古屋市昭和区御器所町字木市29番

TEL:052-735-5627 FAX:052-735-5542

E-mail: nitfair@adm.nitech.ac.jp URL: <https://technofair.web.nitech.ac.jp/>