

# ChatGPTの技術を用いた 固有表現抽出の手法の検討

川村インターナショナル

高木 創志

伊澤 力

# そもそも固有表現抽出とは？

## 固有表現とは？

- 特定の個人、組織、地名など、一意に識別可能な名前を指す言語学の用語。
- 例：人の名前、国、都市、企業、製品名、歴史上の事件など

## 固有表現抽出（認識）（Named Entity Recognition, NER）とは？

- テキストから固有表現を識別・分類すること
- 例：今日、アジア太平洋機械翻訳協会の年次大会にて、川村インターナショナルからの発表があった。
- → 「アジア太平洋機械翻訳協会」、「川村インターナショナル」

# なんのために固有表現抽出を行うのか？

---

## ①固有表現を情報としてまとめたい

→辞書・社名一覧の作成など

## ②テキストマイニング

→SNSやアンケートから人気の観光地・人名などを抽出する、など

## ③個人情報マスキング

→病院のカルテに含まれる個人情報を、外部に開示できるようにマスキングしたい  
など

# 今回の背景

商用機械翻訳エンジンを実際に運用する際に、学習に用いたデータセットから関係がない社名が出現すると、ビジネス上大きな問題が生じる。

→学習データセットの固有表現をマスクする事で、翻訳結果に特定の社名が出現する事を防ぎたい。

→しかし、学習データセットの固有表現をアノテーションしていく作業には膨大な時間がかかり、個別の案件で逐一行う事は現実的ではない。

→ChatGPTのAPIであるgpt-4,gpt3.5-turboを用いてある程度自動的に行えないか？

# 手法

---

約100行の日本語と英語のデータに対して

① **gpt-4にFew-shotのプロンプトを与えたモデル**

② **gpt-3.5-turboをFine-tuning したモデル**

をそれぞれ使用し、固有表現抽出を行った。

# 評価

---

人の手によって固有表現抽出を行ったデータを用いて

- ・ **真陽性**（固有表現を正しく固有表現として抽出した）
- ・ **偽陽性**（固有表現ではない表現を固有表現として抽出した）
- ・ **偽陰性**（固有表現を固有表現として抽出しなかった）
- ・ **真陰性**（固有表現ではない表現を固有表現として抽出しなかった）

をそれぞれ評価した。

# 結果

	gpt-4 (Few-shot)		gpt-3.5-turbo (Fine-tuning)	
	日本語	英語	日本語	英語
真陽性	6	6	7	7
偽陽性	2	2	2	6
偽陰性	0	0	0	0
真陰性	91	91	90	86

gpt-3.5でもgpt-4と遜色ない結果となった。

# 考察

- 固有表現抽出は常に「どこまでを固有表現とするか」という問題を伴う。
  - 商用における実用上は個別の案件に合わせたい場合が多く、汎用モデルよりもある程度チューニングを行いたいシチュエーションが多い。
- API利用料金の面を考慮すると、gpt-4よりもgpt-3.5-turboの方が圧倒的に安い。また、fine-tuningモデルの方が出力が安定するので、現状では固有表現抽出に対してはgpt-3.5-turboのfine-tuningを用いた方がよい。





---

## 今後の展望

- 固有表現をマスキングしたデータを用いた追加学習モデルを作成し、翻訳の精度を評価する。
- マスキングしたデータを用いた追加学習モデルと、置換した追加学習モデルと、汎用モデルで翻訳した後に固有表現をマスキングをするべきかを検討する。