

実世界知識を活用した質問応答システム用の意味解析

自然言語理解に向けて、意味解析技術と実世界知識を繋げましょう。

Connecting Real World Knowledge with Semantic Parsing

ナバーロ ホアン, 宮尾 祐介

どんな研究？

- * 複雑な質問にも正確に答えられるシステムの研究
- * 実世界知識のデータベースを利用するための自然言語の意味理解の研究
- * 現在、何かのエンティティに関連している検索は約50%*

* Pound et al. 2010. Ad-hoc object retrieval in the web of data. In: Proceedings of the 19th International Conference on World Wide Web. pp. 771-780. ACM, New York, NY, USA

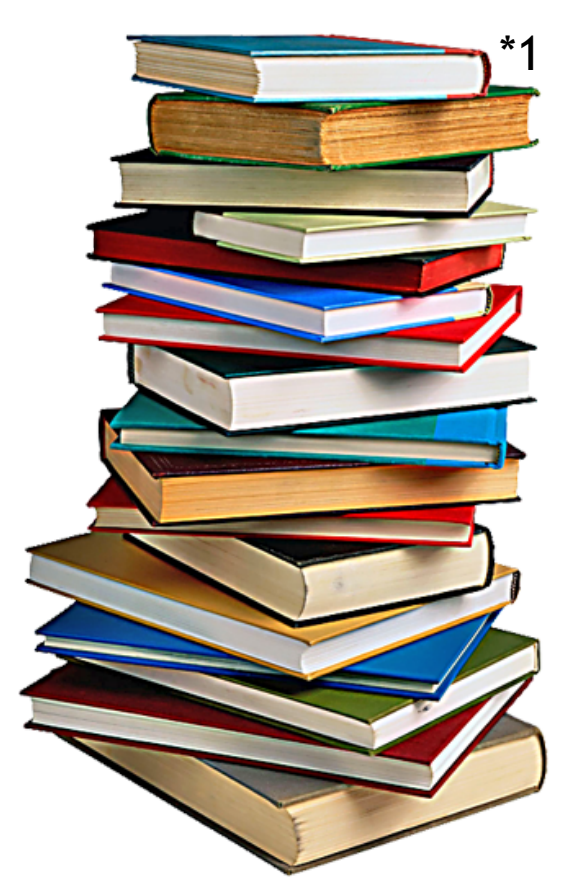
※ (エンティティとは人物、映画、本、歴史的なイベント、などのような固有名詞のことです)

何ができる？

- * 幅広い情報資源の活用が可能になる。
- * コンピュータが人間のように**意味を理解**する。
- * 厳密な解答を与える質問応答システムを実現する。

問題設定

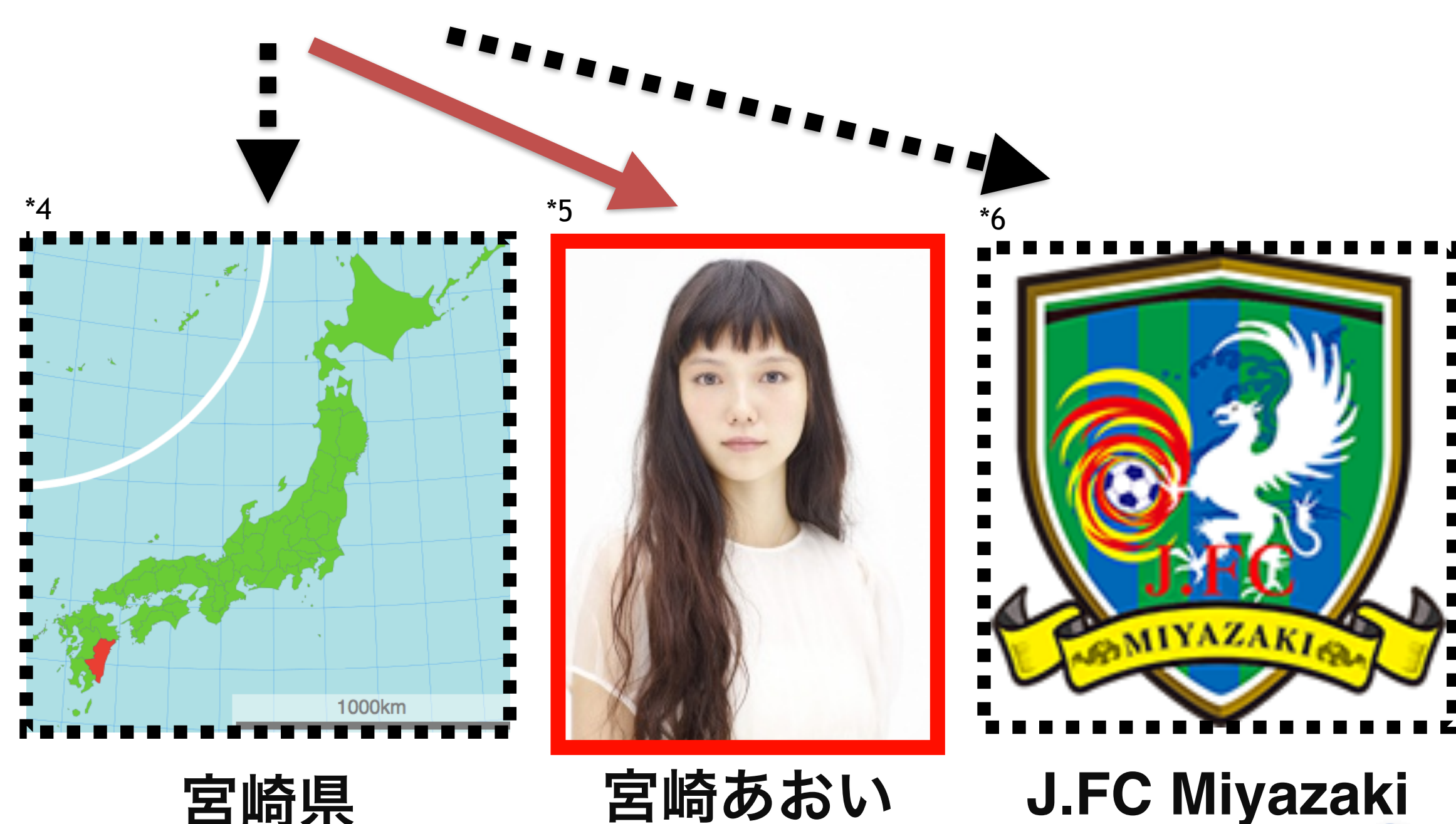
- * 現状、具体的にどこかのドキュメントに明示的に書いていない質問の場合、解答をまとめるのは大変です。(例：地球上で3776メートル以上の山がいくつあるか)
- * いまだに検索エンジンでキーワードを打っていますが、推論や複雑なまとめが必要な質問はキーワードだけでは表せません。
- * 実世界情報を扱うために、データの理解とその情報に基づいた推論が必要です。



研究内容

自然言語文

Q: 宮崎はいくつの賞をもらったか。



意味表現

count($\lambda y. \exists x. \text{person}(x) \wedge \text{name}(x, \text{“宮崎”}) \wedge \text{award}(y) \wedge \text{awarded}(x, y)$)

チャレンジ①: (曖昧性解消)
エンティティ・リンキング

チャレンジ②:
意味表現への変換手法

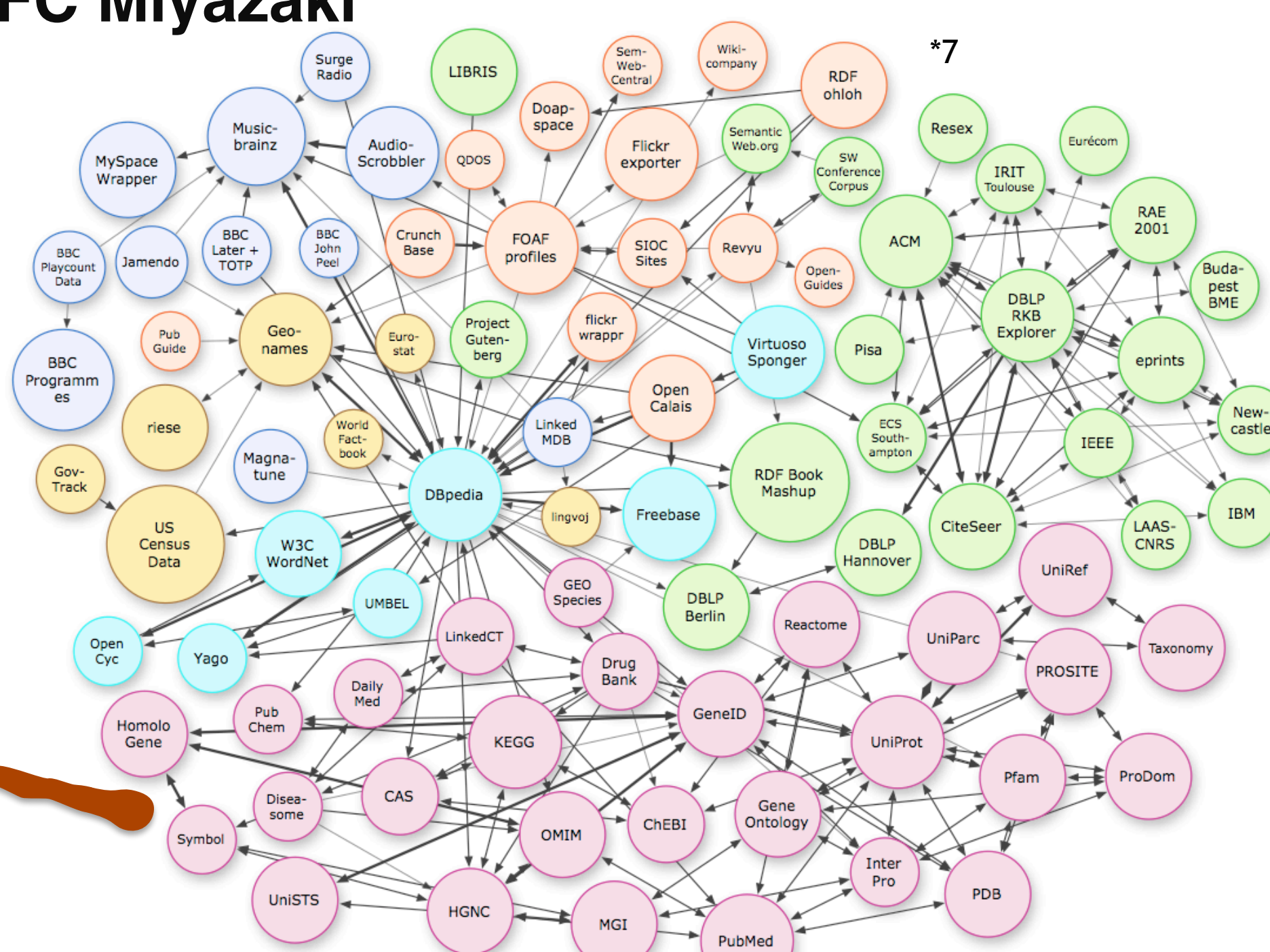
知識ベースの部分グラフ

```
SELECT count(DISTINCT ?award)
WHERE {
  ?person foaf:name “宮崎あおい”@ja.
  ?person <http://dbpedia.org/ontology/award> ?award .
  ?award rdfs:range dbo:Award
}
```

RESULT:
count(
<http://ja.dbpedia/page/第33回日本アカデミー賞>,
<http://ja.dbpedia/page/第35回日本アカデミー賞>,
<http://ja.dbpedia/page/第37回日本アカデミー賞>,
...)

チャレンジ③:
知識ベースの選択

チャレンジ④:
世界知識データの構造



知識ベースの
利用統計データ



チャレンジ⑤:
常識やコンテキストを取得する

*1: <http://clipart-library.com/images/ATbrxyjc.jpg>
*2: <http://consulmatome.com/wp-content/themes/consulmatome/images/qa/shitsumonsha1.png>
*3: <http://consulmatome.com/wp-content/themes/consulmatome/images/qa/shitsumonsha2.png>
*4: https://en.wikipedia.org/wiki/Kanagawa_Prefecture#/media/File:Map_of_Japan_with_highlight_on_14_Kanagawa_prefecture.svg
*5: <http://www.cinemacafe.net/article/img/2017/04/20/48775/278080.html>
*6: <http://jfc-miyazaki.com/wp-content/uploads/2017/05/logo-2-1.png>
*7: http://images.slideplayer.com/27/9127270/slides/slide_18.jpg
*8: <http://cdn1.gbh.com.do/wp-content/uploads/2014/07/statistics.jpg>