

ニュースにおける特定単語の文脈の変遷 ～「ある単語と共に使われている単語」の経年変化を探る～

群馬県立高崎高等学校 高々_物理部_吹部 二年 寺田光里 高田悠希



研究概要 / 目的



本研究は、ニュースの見出しや新聞記事における、「特定の単語の使われる文脈の変遷」を調べるため、複数の単語に対して「その単語の周辺に位置する単語の経年変化」を調査したものである。

数年で世の中が大きく変化する近年。新型コロナの流行、IoTの一般化、ウクライナ情勢など、目まぐるしく変わる社会の中で、

単語がどんな文脈で使われているかが変化しているのではないかと？

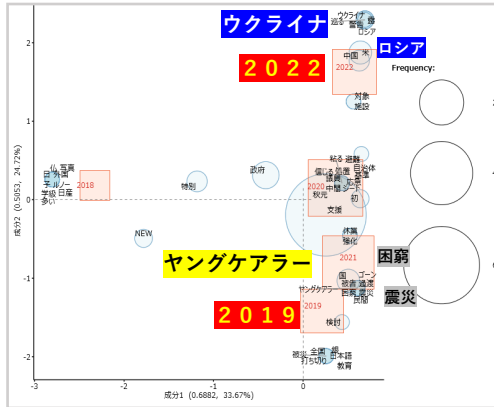
それを調べることで、その単語へに関する人々の関心やイメージ、まつわる社会情勢だけでなく、忘れられた社会課題や、関連する技術の流行の始まりまでもがわかるのではないかと？

と考えた。それらを知ることは、これからの社会研究や社会課題の発見、新製品の開発に至るまで、多くの分野で重要になるはずだ。先行の研究でもメディア記事の内容をテキストマイニングで分析するものはあるが、時間経過で追ってはいなかったり[1]、「終活」など特定の単語のみに注目したもの[2]が多く、汎用的に特定の単語の周辺単語の変化をとれるようにしたものではない。

本研究は、複数の単語に対して時間経過とともに周辺単語がどう変わったかを調べ、そこから資料全体の頻出語を見るだけではわからない、その単語にまつわる**当時の社会情勢・出来事や、かつての人々の関心ごとなどを知ることができると示すものである。**

研究の手法と結果

実験① 「Yahoo!ニュース」の見出し語を抽出、**支援**と共に使われている単語と、年とを対応分析(KH Coderを使用)



2022：ウクライナ支援



使用文脈が変化!!

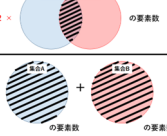
2019：ヤングケアラー支援



実験②-手法 地方紙である「上毛新聞」のデータベースから22年分の全記事を収集
※データ提供：上毛新聞社様

任意の単語を指定し、その単語の周辺単語の年ごとのDice係数を導出、Dice係数の経年変化をグラフ化するプログラムを「Python」上で実装

Dice係数は集合の類似度を測る係数で、2集合の平均要素数と共通要素数の割合。



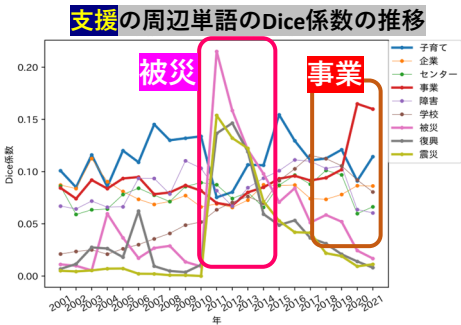
一方の集合だけ要素数が膨大でも大丈夫で、基にした異なる記事や資料同士でもデータ比較ができる →データに汎用性を持たせる

先行事例のほぼないローカル紙からでも汎用的に単語の文脈の変化をグラフ化でき、当時の出来事や情勢などを導けることを示す!

年の経過によって「支援」の語が指す対象が変遷している!!
ヤングケアラーなど…関心が薄れてしまった課題?

実験②-結果・考察 ※全てDice係数の最大値が一定以上のもののみプロットしている

当時の情勢を反映!!



2011年 - 被災 × 支援
東日本大震災(2011.3.11)

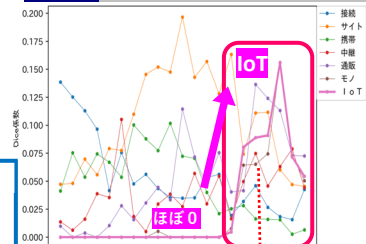
2020年 - 事業 × 支援
COVID-19と緊急事態宣言に伴う事業支援(2020～)

2010頃 - インフルエンザ × 流行
H1N1インフル(豚インフルエンザ)の世界的流行

2014年 - PED × 流行
豚流行性下痢(PED)の世界的パンデミック

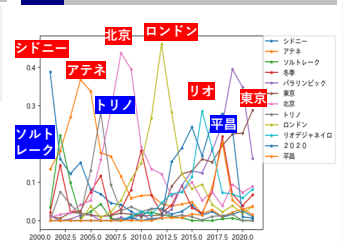
共に「子豚」も急増
養豚が盛んな群馬県の地方新聞ということもあり、かなり顕著

インターネットの周辺単語のDice係数の推移



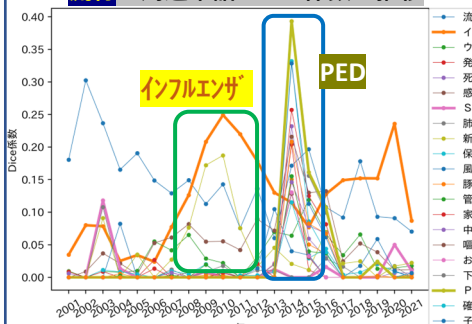
2017年が「IoT元年」だが、その少し前から話題だったとわかる

五輪の周辺単語のDice係数の推移



夏季大会の開催地との共起度が冬季に比べて大きい
→夏季大会の方が話題性が強い?

流行の周辺単語のDice係数の推移



結論と今後の展望

単語の周辺にある語の変遷を調べ、それをグラフ化する汎用的なプログラムを実装できた。また、そのグラフから、**特定の単語にまつわる「当時の情勢」や「人々の関心・話題」**など、全体の頻出語のみではわからないことを知ることができると示せた。この方法は、分野を絞って**詳細に社会情勢や当時の関心事の把握をする**だけでなく、関心が薄れてしまった**現在まで残る課題の再発見**や、**企業やイベント等のイメージ調査**といった幅広い面での活躍が期待される。今後は、**ツール、アプリとして開発**していくことで、より汎用性を高めて誰もが手軽に利用できる手法にしていきたい。

[1]: 樋口, 「現代における全国紙の内容分析の有効性——社会意識の探索はどこまで可能か——」 (2011) https://www.jstage.jst.go.jp/article/jbhm/38/1/38_1_11_pdf-char/ja
[2]: 木村, 安藤 「マス・メディアにおける終活のとらえ方とその変遷—テキストマイニングによる新聞記事の内容分析—」, 横浜国立大学技術マネジメント研究会, 技術マネジメント研究第17号 1-19 (2018)