

Sentence BERT を用いた欧州議会議事録分析による サステナビリティトレンドの重要度評価

松元拓也[†] 林高樹[‡]

慶應義塾大学 大学院経営管理研究科^{†‡}

1. 導入

サステナビリティ（持続可能性）に関する意識が世界的に高まっている傾向にある。これは消費者の商品選択から公的機関の政策決定にまで及んでいる。この傾向を企業経営の観点から見た場合、企業としてサステナビリティを考慮しないことが事業価値の毀損に繋がる環境へと変化していると解釈ができる。言い換えれば、サステナビリティがリスクとして顕在化している時代が訪れている。

このようなリスクに対応するため、比較的用いられるフレームワークがマテリアリティ分析である。マテリアリティ分析では、縦軸に「ステークホルダーにとっての重要度」、横軸に「企業にとっての重要度」を取ったマトリクス内にサステナビリティトレンドをマッピングすることで、企業および社会の両方にとって重要なサステナビリティトレンドを特定することが可能であるとされる。

しかしながらこのフレームワークを実際に用いる上で、トレンドの重要度を測定する統一的な方法は存在しておらず、文献調査やインタビュー調査を組み合わせて行うことが一般的である。結果的に個別企業において、各々の主観に頼った重

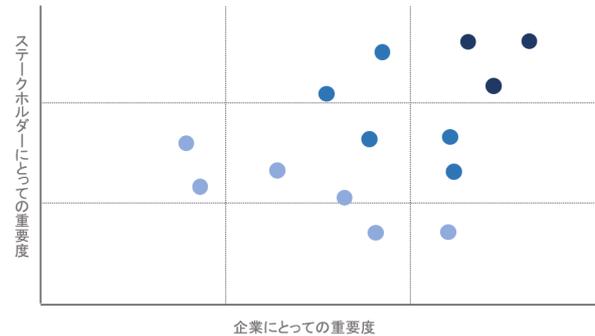


図1. マテリアリティ分析の実施イメージ

要度測定が行われているケースが想定されるだけでなく、企業はその評価プロセスにおいて多くの時間とコストを負担することになる。

実際に、企業を超えた共通課題にもなり得る「ステークホルダーにとっての重要度」すらも、同業界内であるにも関わらず、認識にバラつきが生じている場合がある。

そこで本研究は、主観性をできる限り排除した形で社会全体におけるサステナビリティトレンドを効率的に把握することを目的として、自然言語処理手法を用いた公的機関の議事録分析による重要度評価手法を提案した。

2. 方法

サステナビリティ関連議論をリードする公的機関であると考えられる欧州議会を分析対象とし、Grootendorst [1] によって提案されている Sentence BERT [2] を用いたトピック抽出を行った。

具体的には事前学習かつファインチューニング済みモデルである all-MiniLM-L6-v2 [3] によって議事録センテンスの分散表現を取得した後、次

Assessing the importance of sustainability trends by analyzing the debates of the European Parliament using Sentence BERT.

Takuya Matsumoto[†]
Takaki Hayashi[‡]

Graduate School of Business Administration,
Keio University^{†‡}

元縮約およびクラスタリングを実施した。

このクラスターをひとつのトピックとみなし、さらに全クラスターに対するあるクラスターの占める割合を重要度と定義した。何故ならば、これは一定期間に行われた議論のうち、あるテーマについて割かれた割合と解釈ができるためである。

3. 分析結果

Sentence BERT を用いたトピック抽出を行った結果の例を図2と図3に示す。

トピック「emissions」では、2008年から2009年にかけて高い重要度が検出されており、また近年では再度その重要度が高まりつつある傾向が得られた。一方のトピック「biodiversity」では、徐々にその重要度が高まっている傾向がうかがえるものの、その重要度は「emissions」よりも低いものとなった。

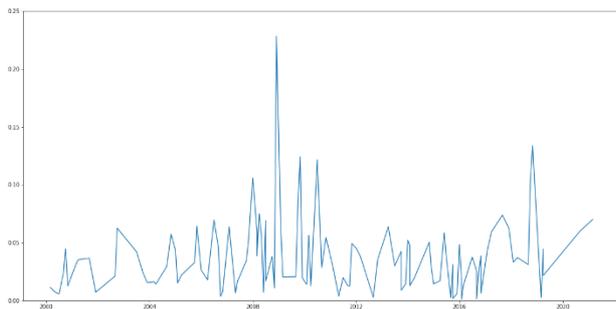


図2. トピック「emissions」の重要度推移

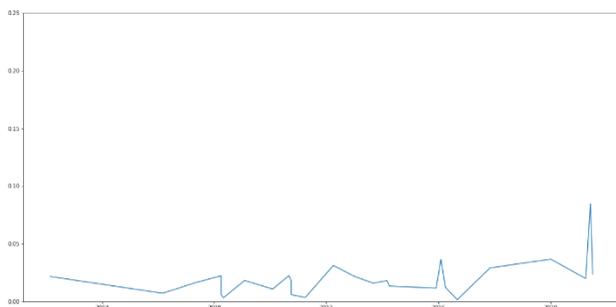


図3. トピック「biodiversity」の重要度推移

4. 考察と課題

文章中に含まれる単語の重要度を評価するための古典的手法である TF-IDF と比較して、本研究

で得られた重要度およびその時間的推移は、より欧州議会における議論内容を反映したものが得られた。そのため本提案の手法が、マテリアリティ分析における従来の評価方法と比較して、より客観的かつ短時間でサステナビリティトレンドの重要度評価が行える可能性が示された。

一方で、Sentence BERT の後継バージョンが提案されている等、自然言語処理分野では新たな試みが日々更新されている状況にあるため、必ずしも本提案で使用した all-MiniLM-L6-v2 で得られる分散表現が最良とは現時点では言い切れない。そのため別の自然言語処理モデルを使用した場合と比較検証を行う等、採用するモデル選択を探索的に最適化していく必要があると考えられる。

また all-MiniLM-L6-v2 をそのまま採用する場合にも、現時点では欧州議会のみを対象とした分析に限られているため、例えば国連や米国議会等へと分析対象を拡大することで、より社会全体における重要度を評価できる可能性があると考えられる。加えて、この方法によって得られる重要度指標の企業経営への具体的な応用プロセスも課題である。

また時系列的にトピックを分析することで、トレンドの中でも特に規制対象となったトレンドに関する特徴分析や、今後の重要度推移予測モデルの構築が期待される。

参考文献

- [1] Maarten Grootendorst.: Interactive Topic Modeling with BERTopic, Towards Data Science, <https://towardsdatascience.com/interactive-topic-modeling-with-bertopic-1ea55e7d73d8>, 2020.
- [2] Nils Reimers., & Iryna Gurevych.: Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, ArXiv e-prints 1908.10084, 2019
- [3] sentence-transformers/all-MiniLM-L6-v2.: Hugging Face, <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2>, (accessed Dev 20, 2020)