

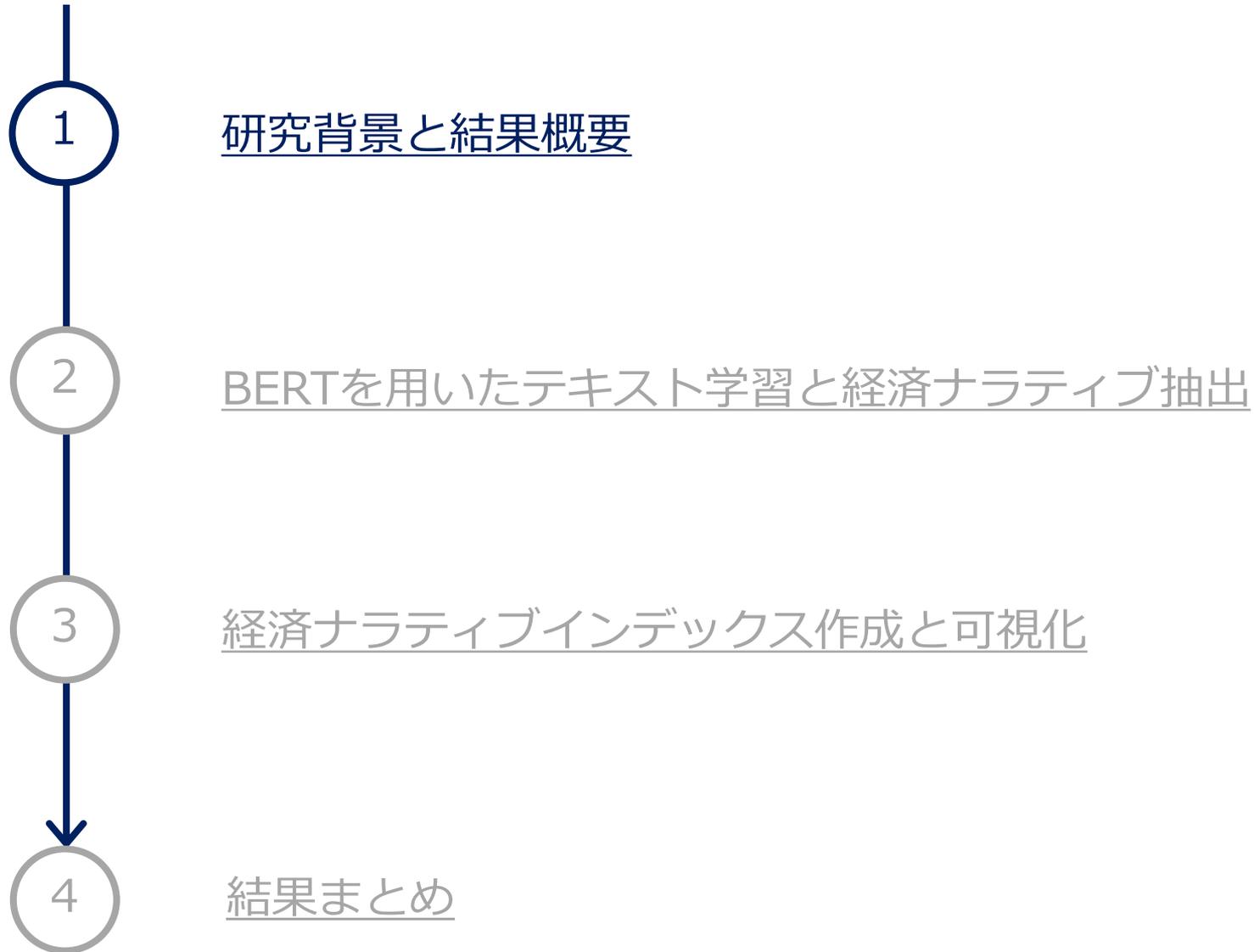
BERTと因果抽出を用いた 気候変動ナラティブの可視化/指数化

2022年11月11日

東京大学大学院工学系研究科
日本銀行金融研究所

坂地 泰紀
金田 規靖

※本資料に示されている意見は、筆者ら個人に属し、日本銀行の公式見解を示すものではない。
また、ありうべき誤りはすべて筆者個人に属する。本資料はファイナンス・ワークショップ参加者限り。



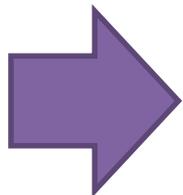
研究背景と概要

- 気候変動がマクロ経済・物価に与える影響に関する経済研究は様々な分析手法やデータを用いて行われている。
 - 政府、企業、金融機関など様々な経済主体が、気候変動リスクへの対応に直面
 - 企業の気候変動対応が株価・業績パフォーマンスに影響を与えるとの報告もみられている
- 近年、自然言語処理分野において、Deep Learning(DL)ベース言語モデルのBERT (Bidirectional Encoder Representations from Transformers) が注目されている。当モデルはテキストの文構造や文脈を学習できる点が特長。
 - 経済ファイナンス分野では、テキスト分析が増加。中銀では、ナウキャストイングやセンチメント指標作成などでの利用が拡大。ただし、統計的言語処理等の非DLモデルが主流。



BERTと因果抽出を用いて、「気候変動」に関する経済ナラティブ (原因＝結果のつながり)を指数化/可視化する手法を提案

※因果抽出 (Sakaji *et al.* [2008]等)を利用



従来のテキスト分析とは異なり、テキストの原因＝結果関係を解析した経済ナラティブを抽出できる

テキスト学習
因果抽出

- ①BERTで、日経記事データを学習
- ②因果抽出で、経済ナラティブを解析

指数化
可視化

- ③経済ナラティブインデックスの作成
- ④経済ナラティブネットワーク関連図の作成

経済的解釈

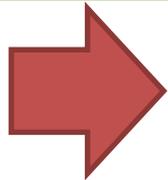
- ⑤経済ナラティブの分析

今回は気候変動ナラティブを抽出し、分析する

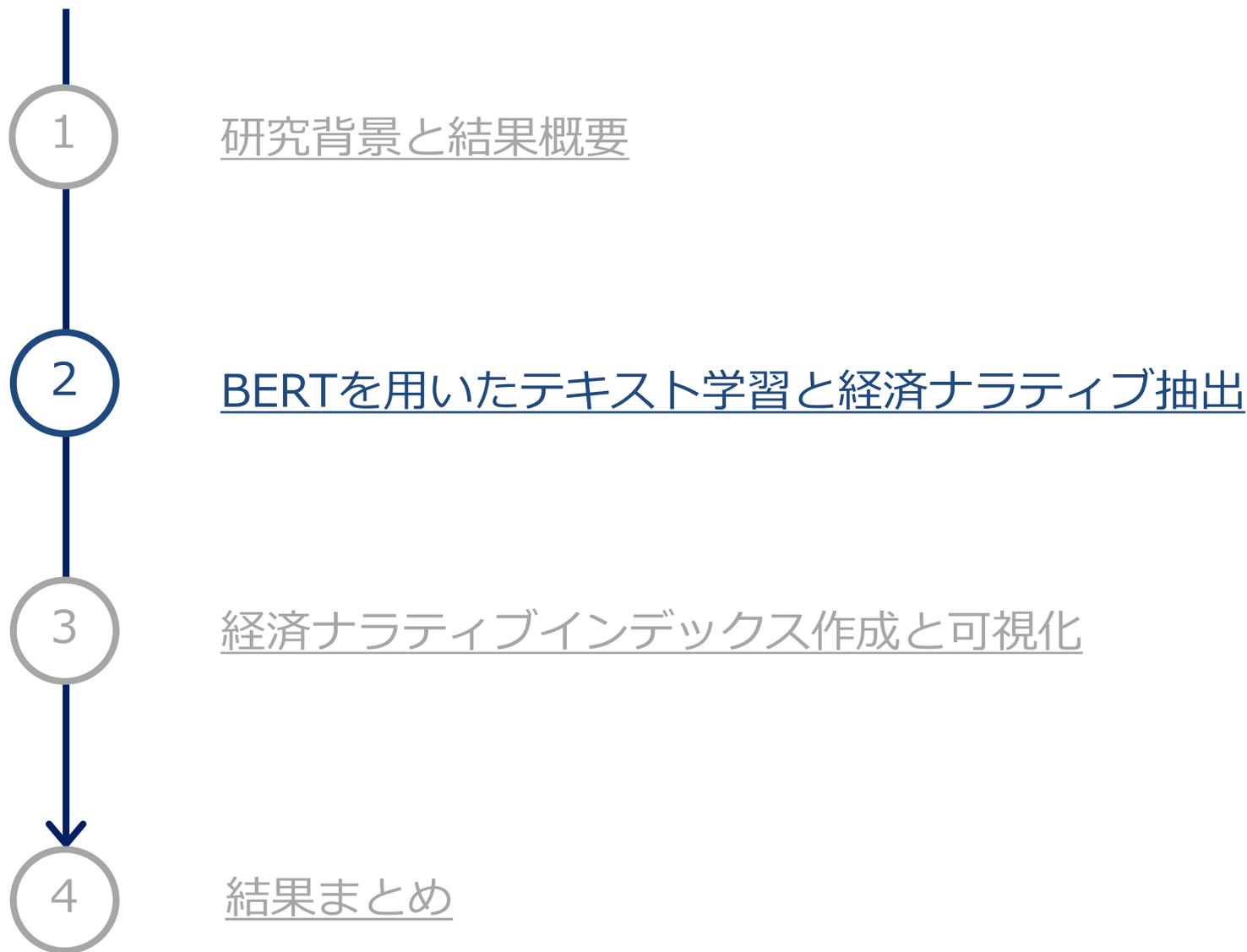
結果概要

- ①BERTを用いることで、文脈を考慮して日経記事データを分類
- ②因果抽出とチェーン構築を用いることで、任意のトピックに関する経済ナラティブ（原因事象と経済的帰結）の出現頻度やつながりの強さを抽出
- ③「ある時点のニュース（経済的帰結）が、過去のニュース（原因事象）からどの程度影響を受けているか」を指数化する方法を提案し、「経済ナラティブインデックス」を作成
- ④経済ナラティブをネットワーク関連図として可視化
- ⑤本研究の気候変動に関する経済ナラティブは、気候変動対応の議論が進捗し、環境制度・規制として具体化される中、認識変化や行動変化が始まっていることを示唆

Shiller [2019]は 人々の物語（経済ナラティブ）が経済を動かすと主張している。例えば、ある物語は根拠なき熱狂となって人々の信念や行動を変えて、マクロ経済を動かした（不動産バブル、ビットコイン、AIによる雇用喪失への不安）という仮説を指摘。



我々の研究は、この仮説に対する回答の一つ



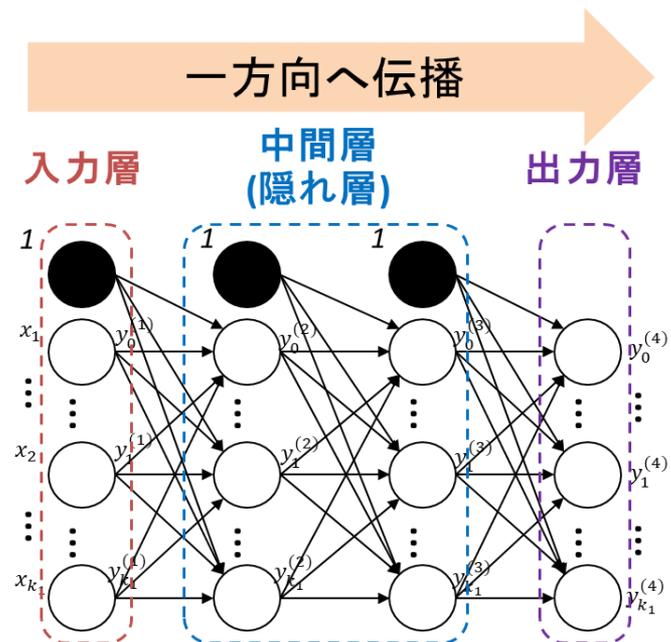
BERTの概説 (1/6)

■ BERTを理解するためには、以下の要素を把握する必要あり

- ニューラルネットワークモデル (MLP、RNN)
- Word2vec (分散表現 : Word Embedding)
- Attention Mechanism (注意機構)
- Transformer
- ELMo

■ MLP (Multilayer Perceptron)

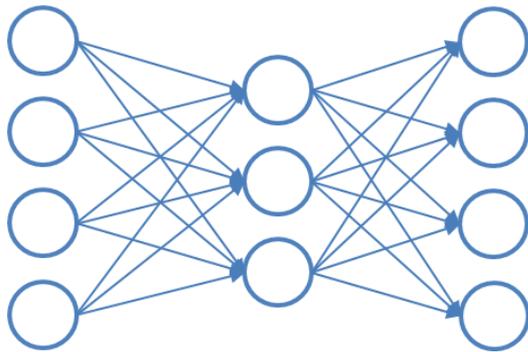
- 多くの層を持ったニューラルネットワークモデル
 - 階層型ニューラルネットワーク
- 誤差逆伝播法により、各層の間にある重みの値を繰り返し更新することで、学習を行う



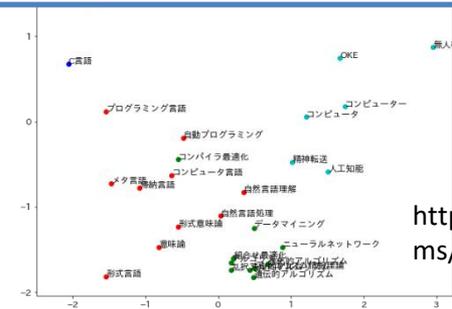
BERTの概説 (2/6)

■ Word2vec (分散表現 : Word Embedding)

- 単語の意味や文法を捕らえるために、単語をベクトル表現化して次元を圧縮したもの
 - ここでの意味とは、類似した使われ方をする語は同様の意味を持つという仮定から定義される
- ニューラルネットワーク(MLP)で学習される



学習された単語ベクトルは類似した単語を調べるなどに利用可能

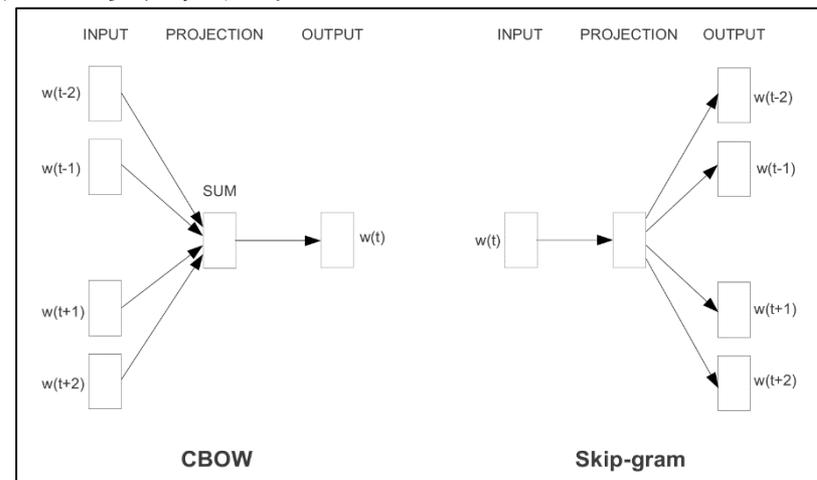


<https://qiita.com/hima2b4/items/0123158097c7840f7b4e>

学習アーキテクチャー:

Mikolov et al. [2013]

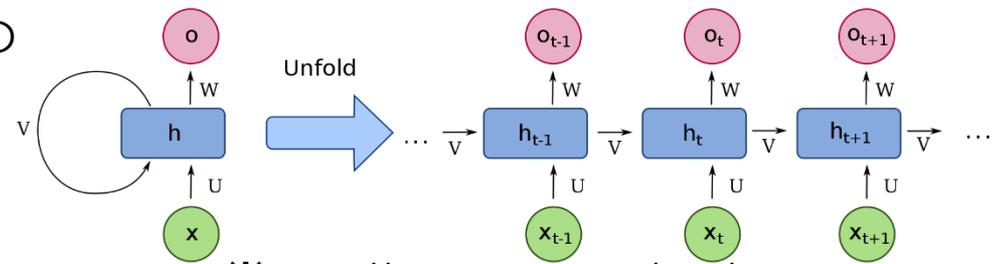
- ただし、入力する単語は独立して扱う (Bag of Words)
- 次に説明するRNN系列の入力に用いると性能が向上することが知られている



BERTの概説 (3/6)

■ RNN (Recurrent Neural Network)

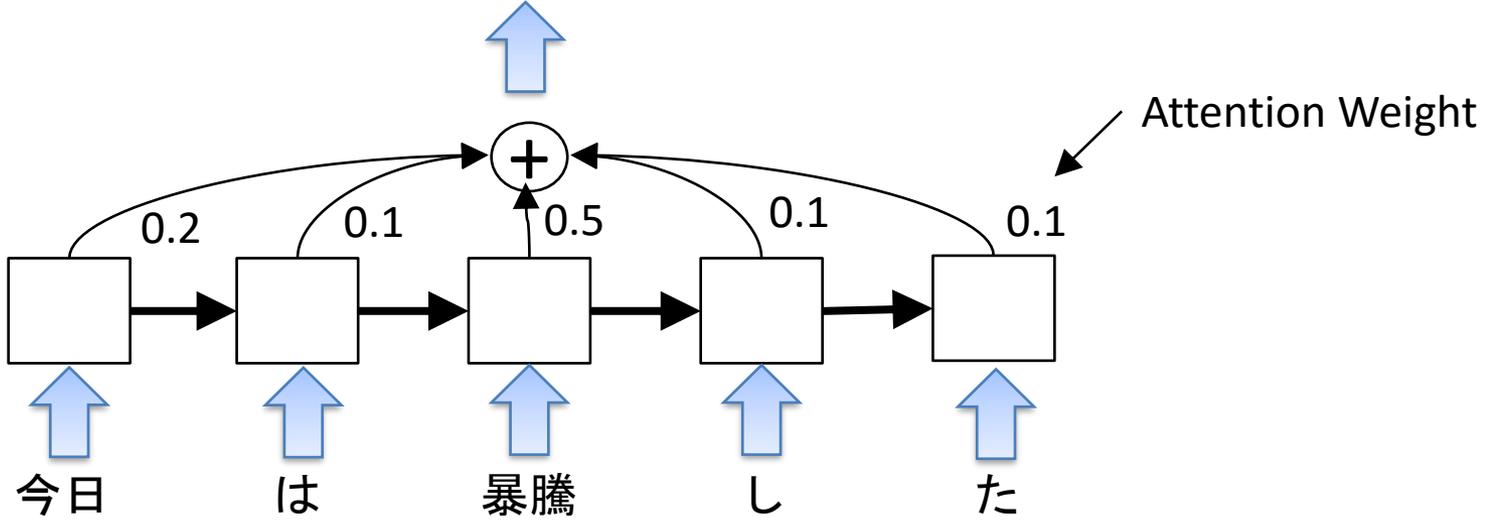
- 自己回帰型の構造をもつニューラルネットワークの総称
 - 系列データに対して、非常に有効
 - 分散表現との相性が○



※ https://en.wikipedia.org/wiki/Recurrent_neural_network

■ Attention Mechanism

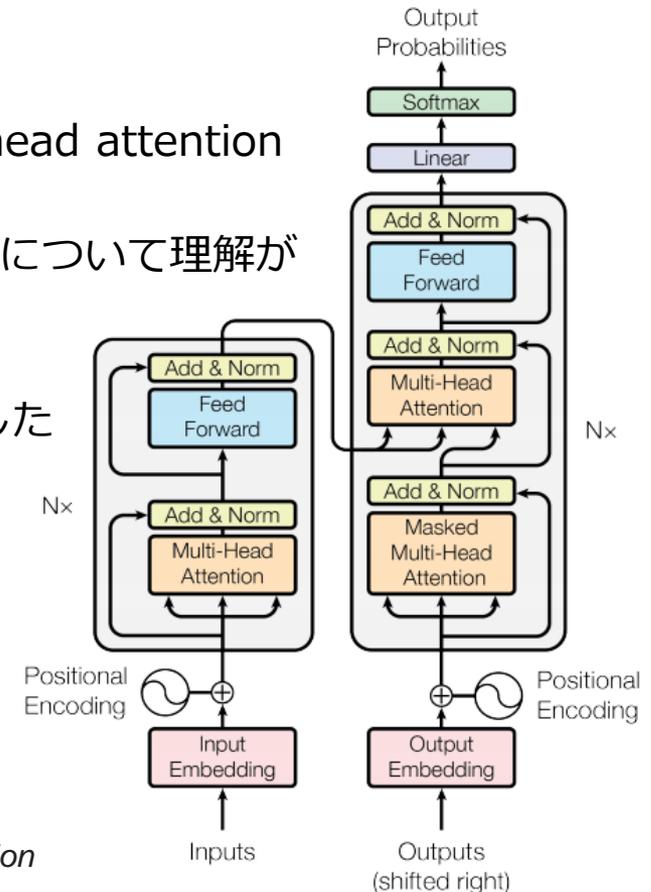
- RNN (Recurrent Neural Network) 系のニューラルモデルから出力を得る際に、どの時点の出力を注視すべきか (重みをおくか) を自動的に学習させる機構



BERTの概説 (4/6)

Transformer

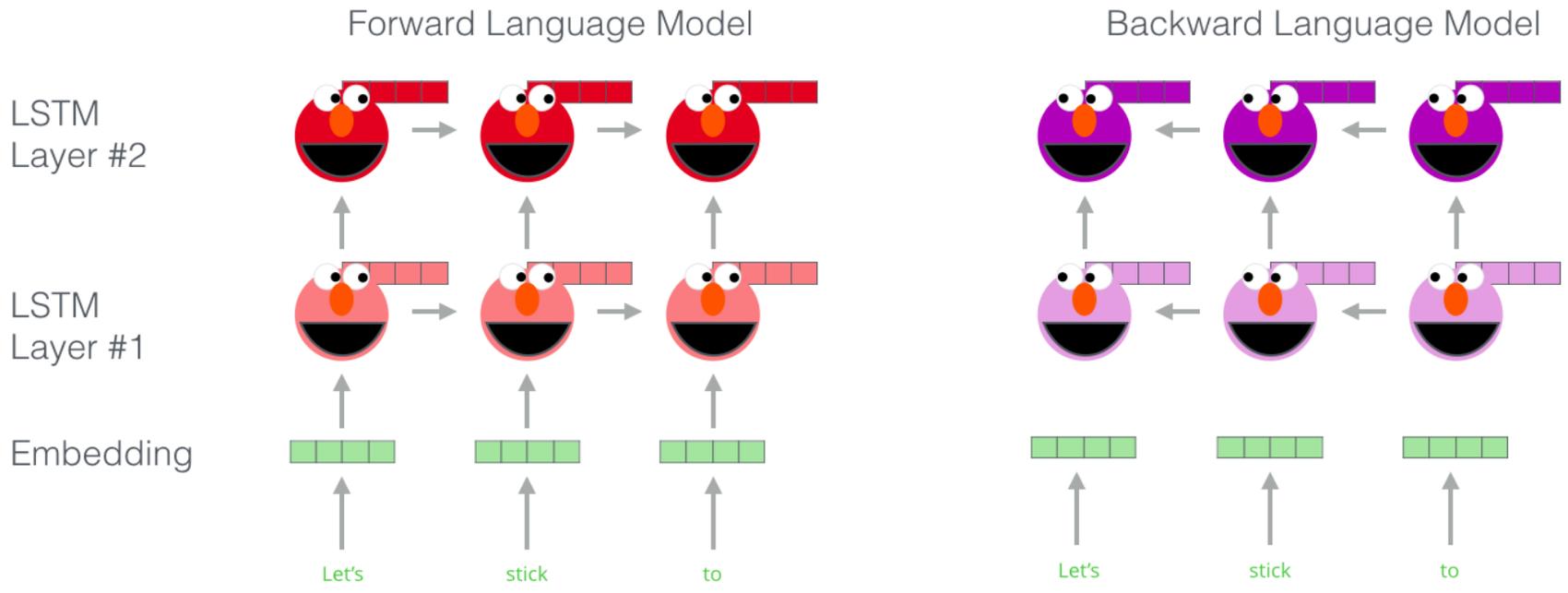
- RNN系ニューラルモデルを並列で動かせるようにした新たなモデル
 - Attention Mechanismを大量に組み込むことで、RNN系ニューラルモデルの再現が可能に
 - Scaled dot product attention + Multi-head attention
 - Attention Mechanism の言語処理への有用性について理解が深まった研究
 - Transformerは実験的試行の結果、精度を示した組み合わせモデル



BERTの概説 (5/6)

ELMo (Embeddings from Language Model) [Peters et al. 2018]

- Word2vecの学習時、周辺単語をbag of wordsで学習していたことから、もし、文脈情報も加味した学習ができれば、良い言語モデルが作れるのでは、という発想から開発
- word2vecは、ある単語について、周りに出現する単語から、その単語を予測するモデルを構築することで、単語ベクトルを作り出すモデル



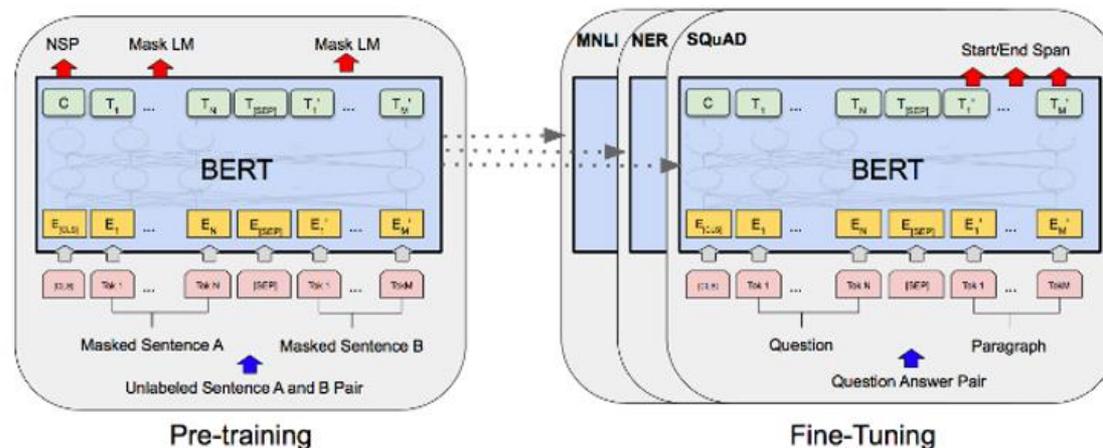
図は<https://jalamar.github.io/illustrated-bert/>より抜粋。

BERTの概説 (6/6)

■ BERT (Bidirectional Encoder Representations from Transformers)

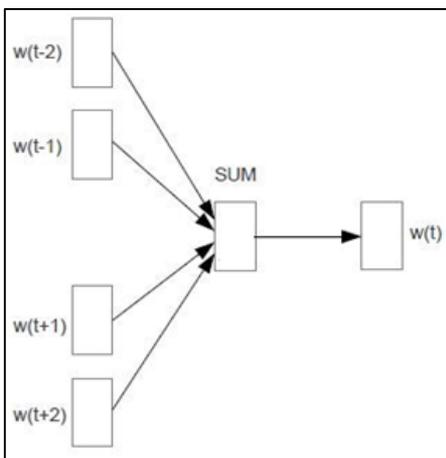
- Transformerを双方向に組み合わせた汎用言語モデル
- Fine-Tuningを行い、タスクに適用させることを前提にしたモデル
- next sentence predictionとmasked word predictionの二つの問題を解くことで学習
- 数多くのモデルでSOTA (State of the Art) を達成

(出所) Devlin *et al.* [2018]

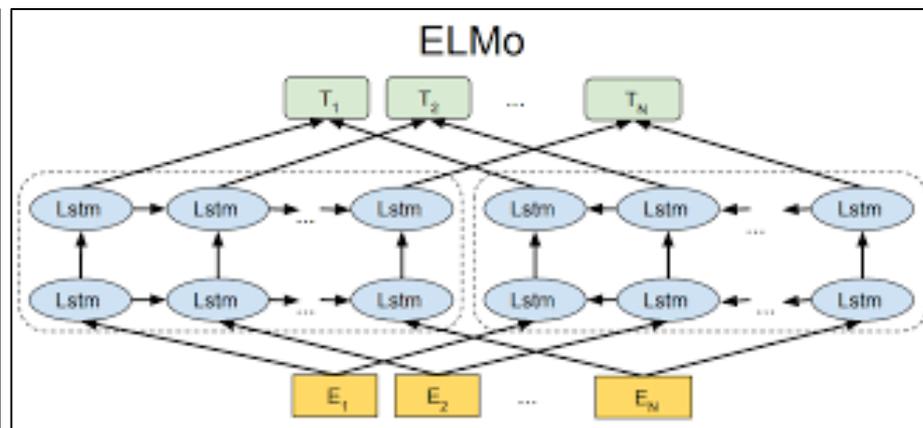


※ ELMoモデルから発想を得たため、BERTと名付けられた

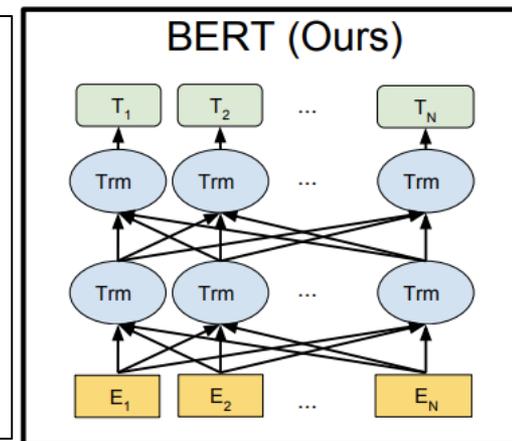
言語モデルの進化



Mikolov *et al.* [2013]



Peters *et al.* [2018]



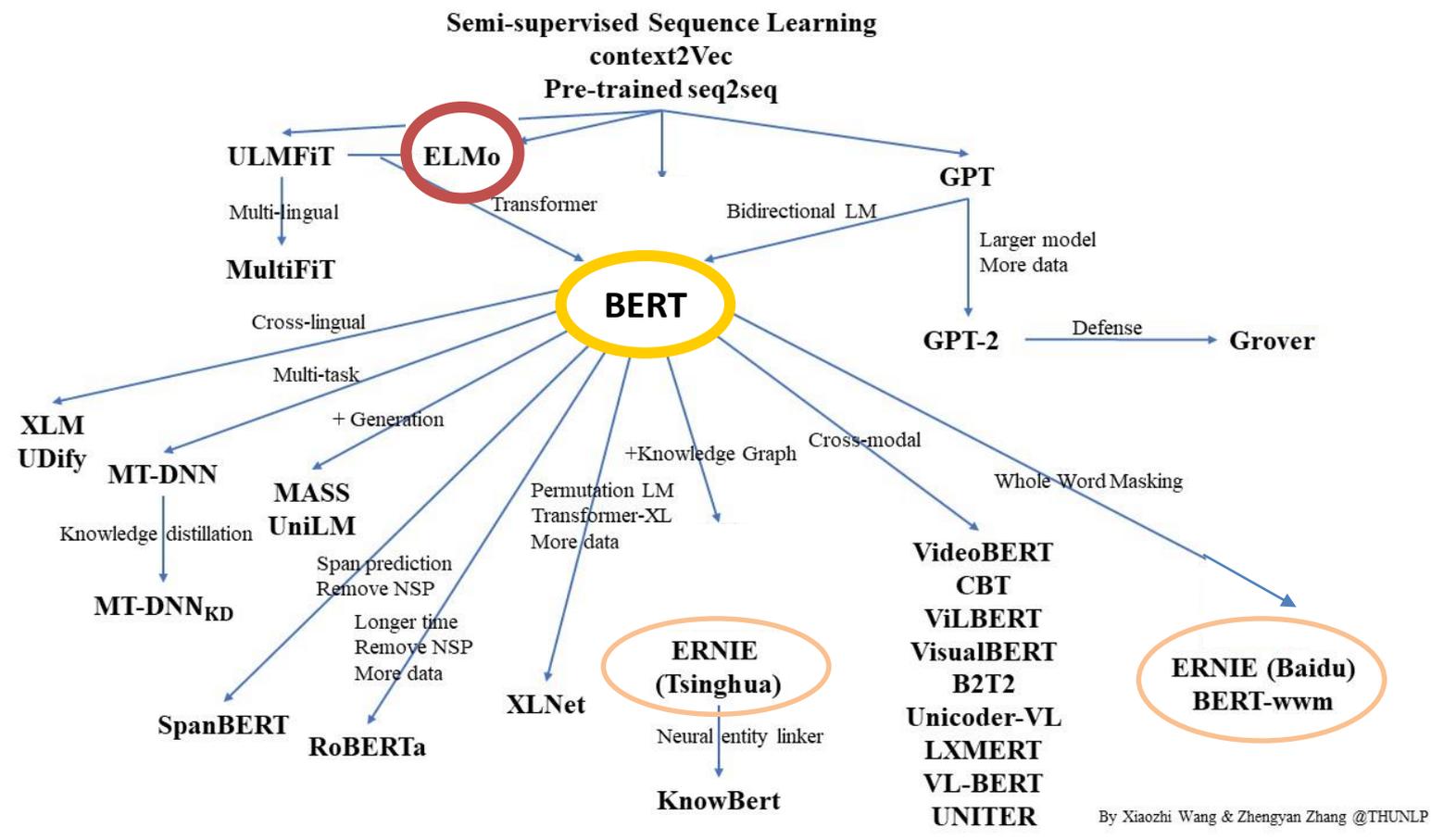
Devlin *et al.* [2018]

word2vec

ELMo

BERT

まだまだ増えるTransformer系



By Xiaozhi Wang & Zhengyan Zhang @THUNLP

まだまだ発展中

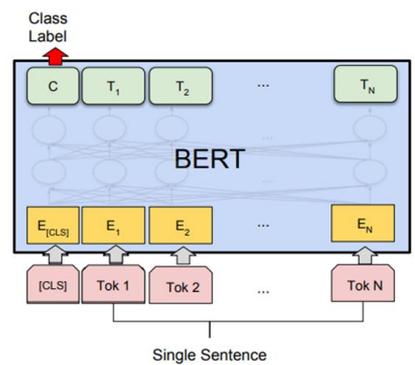
図は<https://github.com/thunlp/PLMpapers>を一部改変。

テキストからの因果抽出

■ Step.1 因果判定 [坂地 *et al.* 2011]

- 機械学習を用いて入力された文に因果関係が含まれているか否か（存在の有無）を判定

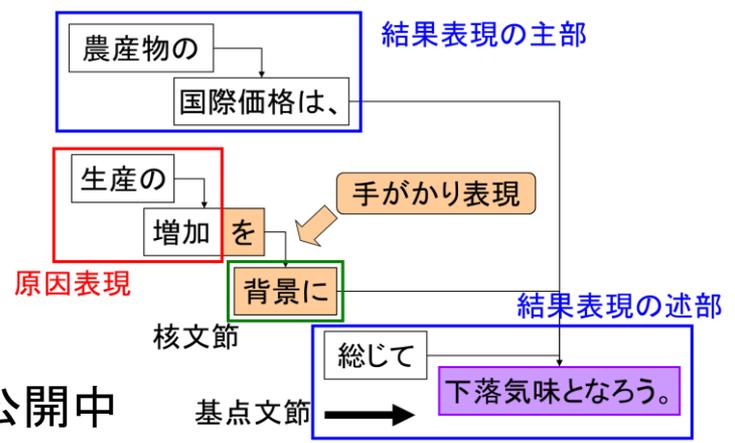
- 今回はBERTを利用



■ Step.2 因果抽出 [Sakaji *et al.* 2008]

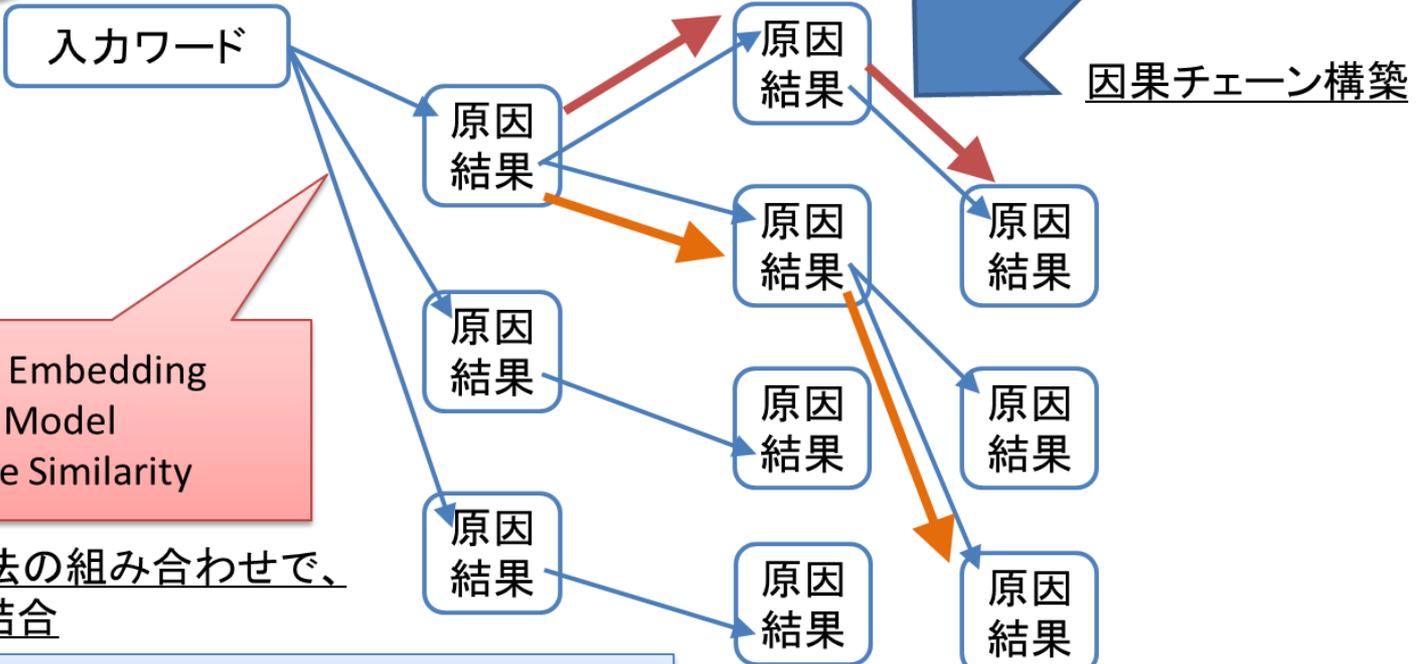
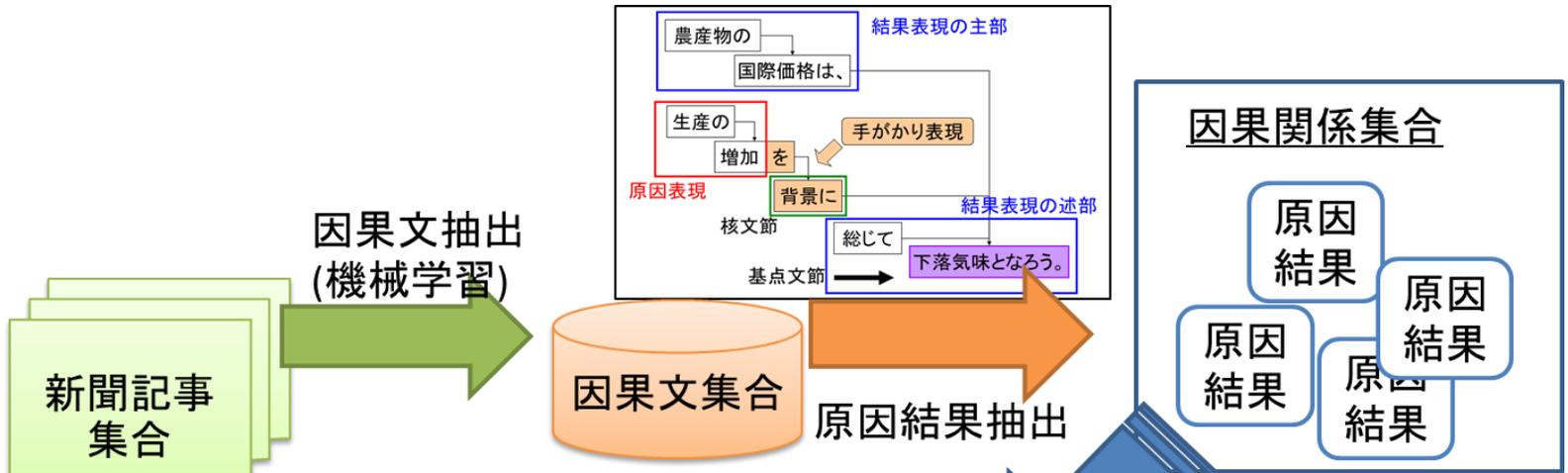
- 因果関係が含まれている因果文から原因・結果表現を抽出

- 係り受け解析に基づく構文パターンを利用



プログラムはGitHubで公開中

抽出した因果情報を用いた因果チェーン構築



- Word Embedding
- Topic Model
- Cosine Similarity

上記手法の組み合わせで、因果を結合

今回は、Word EmbeddingにBERTを利用

因果チェーン検索における事例

- 東京大学和泉研究室で公開している因果チェーン検索システムを用いてみると、



オリンピックから分譲マンションの価格上昇というつながりを見つけることができる。

上記例のように物語を捉えることができることから、

因果ネットワークを
経済ナラティブとして扱うことが可能

※住友不動産の決算短信(2020年3月)には、「不動産販売事業部門において6期連続で過去最高更新」と記載

英語BERTを用いたファイナンス分野の先行研究

- ニュース記事、企業決算などを用いてファインチューニングした各種BERTが提案され、高精度を示している（FinBERT<Liu et al.[2020]>、ESGBERT<Mehra et al.[2022]>など）。
- BERTを用いたファイナンス研究としては、①テキスト分類・キーワード抽出、②BERTの出力結果からニュースインデックスを作成し、金融価格予測を行うなどの事例がみられている。

	先行研究	概要
①	Bingler et al. [2022]	企業の気候変動リスク開示文書を分析するためにClimate BERTを構築（TCFDなど気候変動の公式文章で学習）し、企業TCFD賛同前後の開示情報の実態を分析した結果、 <u>大半の企業は重要度の低い項目の情報開示を増加させているが、経営戦略などの重要な情報開示はさほど進んでいないことが示唆。</u>
②	Kölbel et al. [2021]	移行リスク・物理的リスクを判定するBERTを構築し、 <u>企業の気候変動リスクの開示度を指数化。企業のCDSスプレッドとの関係を統計的に検証したところ、パリ協定後、移行リスク開示進捗は、CDSスプレッド上昇（リスクの織り込みが進む）、物理的リスクの開示はスプレッド低下（将来の不確実性が低下する）につながる</u> ことが分かった。 また、 <u>トランプ大統領当選時は、移行リスクへのエクスポージャーの大きい企業のCDSスプレッドが低下していた（移行リスクの低下を示唆）。</u>
	Sonkiya et al. [2021]	金融BERT（Araci [2021]）を用いて、金融情報サイトのニュースや書き込みから <u>市場センチメント指標を作成。GANモデルを用いてApple株価の予測を行うと、ARIMAやその他のDLモデルを上回る予測精度を示した。</u>
	Mittal et al. [2022]	BERTを用いて、ニュース記事を分析し、Dow Jonesにおける <u>センチメント指数を作成。株価騰落とセンチメント指数の一致度は91.12%を達成し、センチメント指数を組み込んで価格予測を行うと、予測誤差（MSE）は改善。</u>

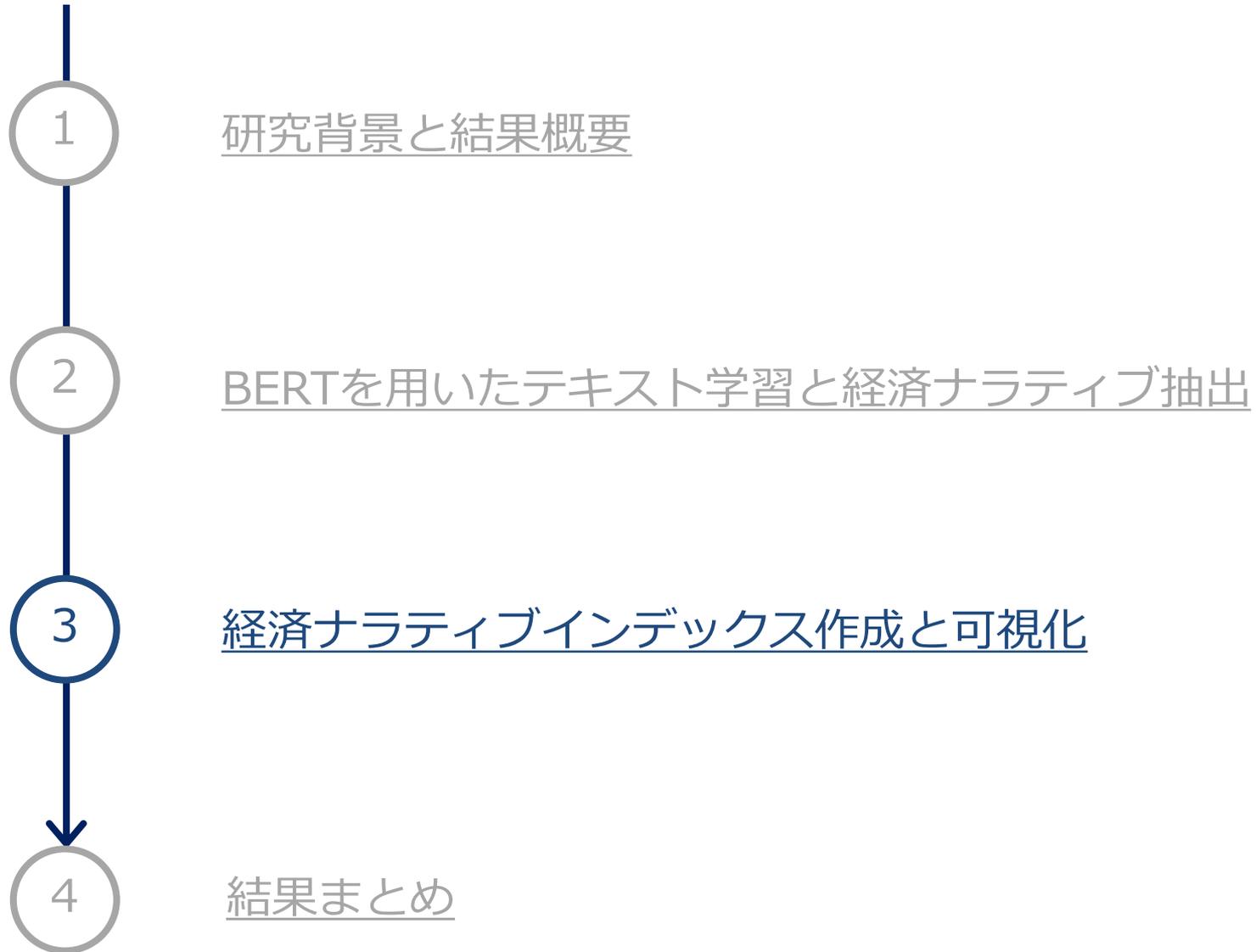
日本語BERTを用いたファイナンス分野の先行研究

- 公開日本語BERTは、東北大乾・鈴木研の日本語BERTや東大・和泉研の金融BERTなどがある。
- 日本語BERTを用いた研究をみると、①ニュースや企業決算などテキスト分類・情報抽出といった事例がみられはじめています。

	先行研究	概要
①	土橋・中田 [2021]	上場企業の有価証券報告書における経営方針項目及び事業等のリスク項目からESG関連文を抽出し、ESG 情報開示の動向を抽出し、可視化。
	前原・久々宇・長部 [2021]	特許文書から脱炭素技術であるか否かを判定する分類モデルを作成し、経産省の産業分類（「エネルギー関連」「輸送・製造関連」「家庭・オフィス関連」）への紐づけを行った。また、各産業における企業の特許出願状況をスコア化。
	若本他 [2022]	企業の統合報告書のESG情報をGRIスタンダードのESG項目別に判定するようモデルを作成したところ従来モデルより高い精度を示した。なお、統一フォーマットのデータは大容量データの効率的なモデル化に有益との指摘。
	毛利他 [2022]	大規模記事データから脱炭素関連トピックの判定モデルを作成し、トピックの出現数を国別にスコア化。日本はCO2回収・貯蔵などの次世代技術が他国より進展している（頻出）が、環境規制が他国より遅れている（僅少）ことが示唆された。

- 一方、②ニュース・インデックスを作成し、金融指標の検証や価格予測を行うといった研究は皆無。また、経済ナラティブを考慮し、ニュース・インデックスを作成した事例は我々の知る限り、存在しない。

⇒ BERTと因果抽出を組み合わせることで、経済ナラティブを抽出できる（①の方向性）。経済金融データを用いた実証研究（②の方向性）にも対応できるよう、抽出したニュース因果を指数化する。また、気候変動経済ナラティブをネットワーク関連図として可視化する。



日経記事データ

【日本経済新聞本紙（朝刊、平日のみ）】

- 期間：2000年1月～2021年11月
- 約137万8千記事（うち気候変動関連は約1万7千記事＊）

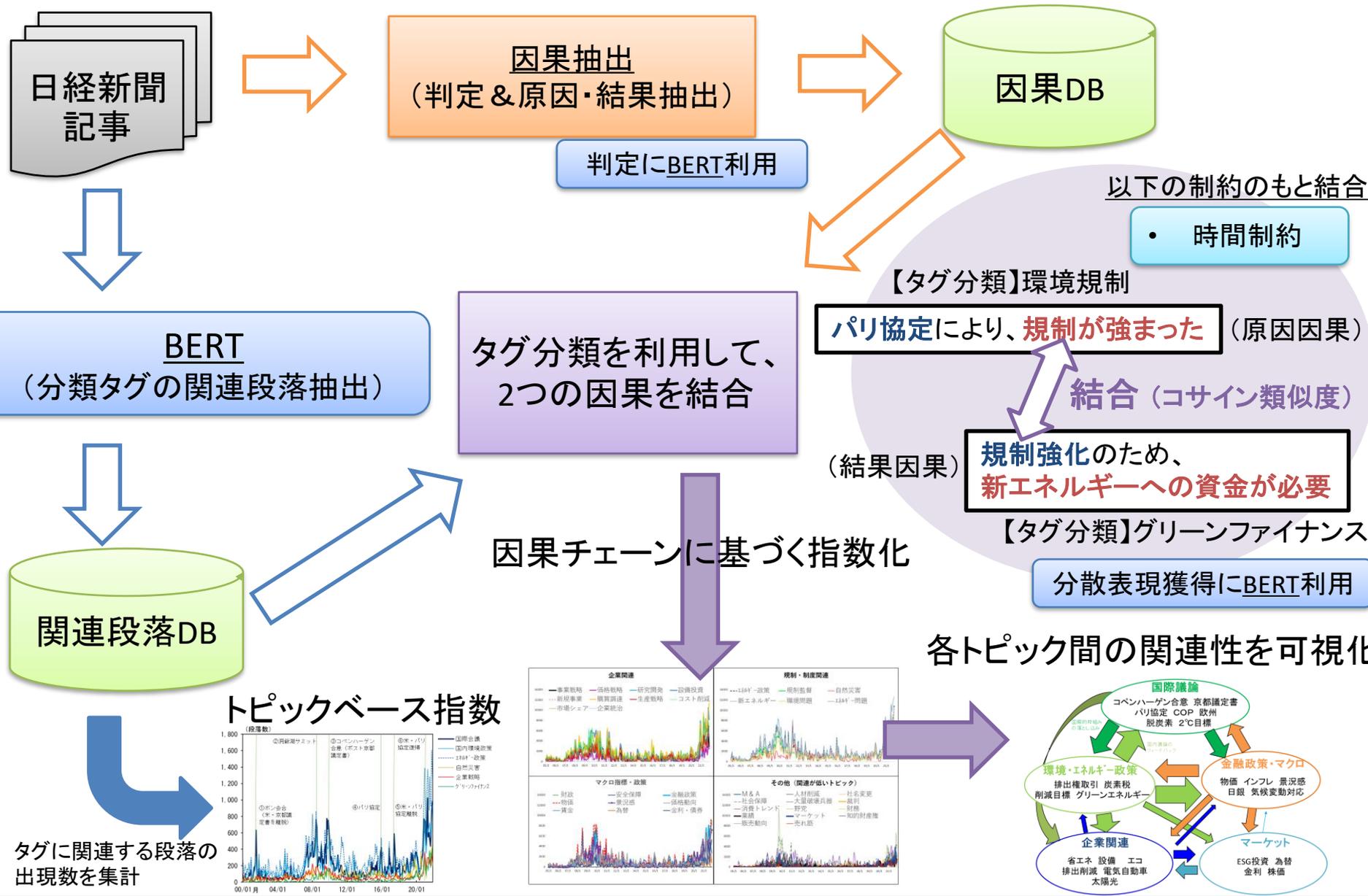
＊「気候変動」「温暖化」「温室効果」のいずれかが一単語以上含まれる記事。

- 定型フォーマットで統一されたデータ（記事ID、記事日付、分類タグ<189種類、記事のトピック>、記事本文など）。

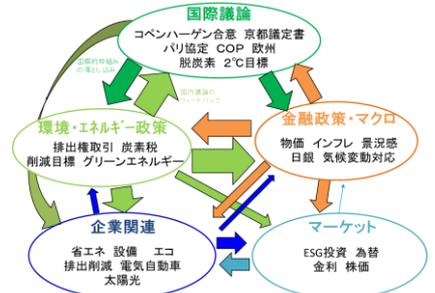
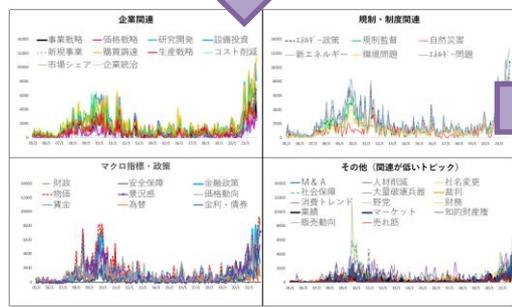
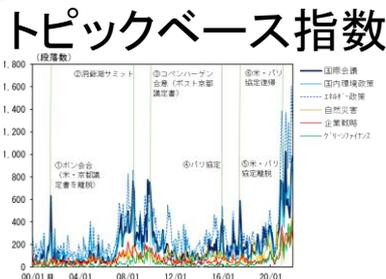
■ 今回は下記の40トピック（分類タグ）を分析対象とする。

【企業】		【政治】	【経済】	【社会】
①新規事業進出	⑪企業統治	⑲野党	⑳金融政策	㉑大量破壊兵器
②事業・企業の買収	⑫人材削減	㉒規制監督	㉓物価	㉔自然災害
③事業戦略	⑬社名変更	㉓国家財政	㉔景況感	㉕裁判
④価格戦略	⑭賃金	㉔IHLG-政策	㉕マーケット	㉖IHLG-問題
⑤生産戦略	⑮財務	㉕安全保障	㉖為替	㉗環境問題
⑥コスト削減	⑯業績	㉖首脳会議	㉗金利・債券	㉘消費トレンド
⑦購買・調達	⑰販売動向	㉗社会保障	㉘新エネルギー	
⑧知的財産権	⑱価格動向			
⑨研究開発	⑲市場シェア			
⑩設備投資	⑳売れ筋			

経済ナラティブインデックス生成過程の概要図

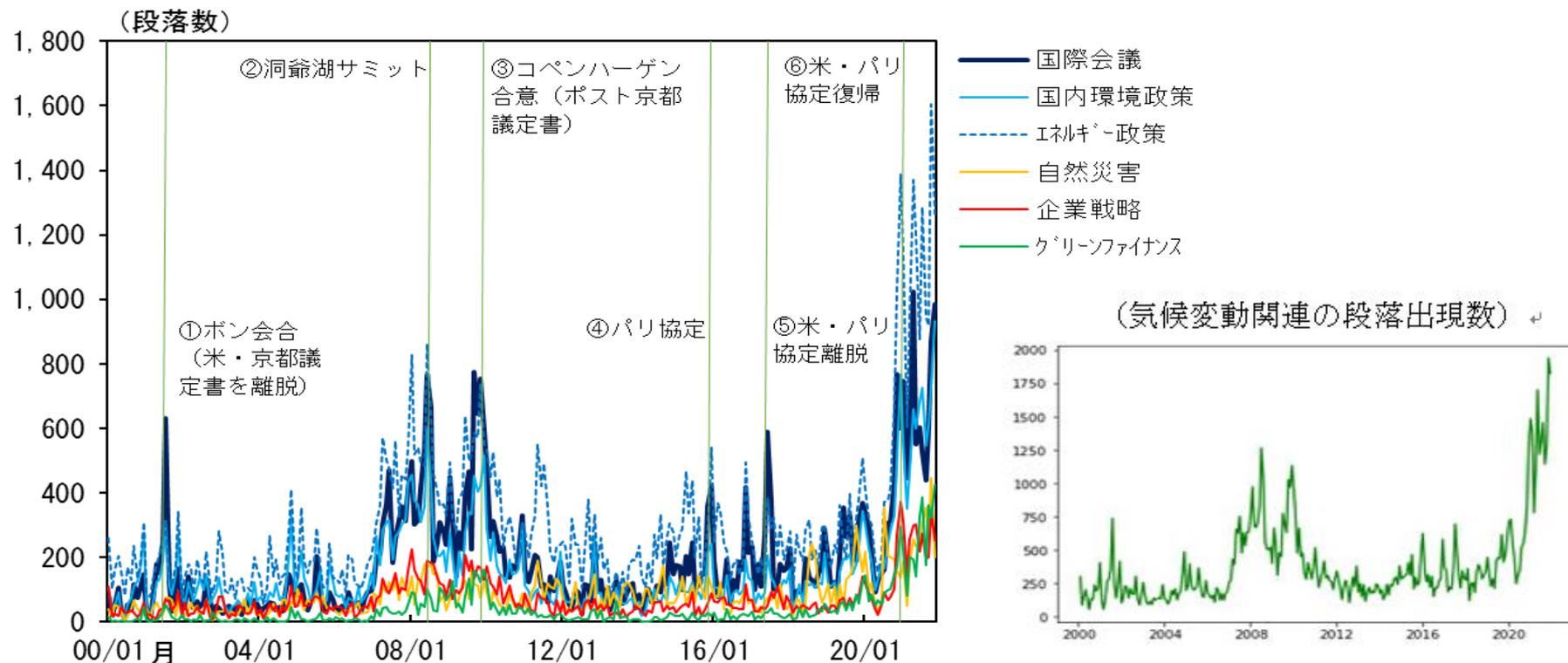


タグに関連する段落の出現数を集計



基礎的分析（トピックベース指数）

○BERTを用いて、気候変動関連記事（段落単位）でのタグ出現数を単純集計。2008-10年頃（洞爺湖サミット、コペンハーゲン合意）や2015年（パリ協定）、米国のパリ協定離脱・復帰などで増加。足もとにかけては、企業戦略などのミクロ的な動きを含め、多くのトピックが急増を示している。



⇒ トピック間には高い相関（自然災害とその他トピックの相関はやや低い）。

⇒ 気候変動のサブトピック同士は共起していることを示唆。

基礎的分析（トピックベース指数）

- 気候変動ニュースのトピックは、時期により移り変わっている様子が窺われる。
 - ・ 2008-10年頃は、**ポスト京都議定書に関する国際議論、国内環境規制・政策議論に関するトピック**が大宗。
 - ・ 2018年以降は、企業戦略やグリーンファイナンスなど**ミクロな企業行動、与信行動に関するトピック**が目立ち始める。
 - ただし、**経済ナラティブがトピック間でどのように連関しているかは、前スライドのトピックベース指数からは明らかではない。**
- ⇒ 「因果抽出」を組み合わせることで、記事データの因果性を手がかりに、気候変動における各トピックの経済ナラティブを抽出できる。これらを指数化することで、トピック間のつながりの強さを定量評価することができる。
- ⇒ 様々なトピックで総当たりに推計したうえで、重要な経済ナラティブを探索し、指数化やネットワーク連関図として可視化する。

経済ナラティブインデックスの作成方法

- 日経記事の全段落があるトピックに属するか否かを分類し、関連文を抽出する。

- ①日経記事データ・分類タグを用いて、金融経済関連の各トピックの教師データを作成する。
 - タグが2つ以下の記事から学習するトピックを含む記事を正例、含まない記事を負例として抽出する（タグが3つ以上の記事は複数トピックが交じった純度の低い記事とみなし使用しない）。
- ②教師データを用いて、BERTをファイン・チューニングし、分類モデルを構築する。
- ③トピック分類モデルを用いて、集計対象記事の各段落がトピックに属するか否かを判定し、各トピックの関連段落を抽出する。

- 因果抽出の結果を用いて、2トピックの因果関係を結合し、経済ナラティブを抽出する。

- ④因果抽出を用いて、各記事における原因＝結果の論理関係を表す表現対を抽出する。
- ⑤各トピックの関連段落（上記③）と因果抽出の結果（上記④）を組み合わせ、最小限の仮定（時制：過去ニュース⇒将来のニュースに波及する）を置き、トピック間の経済ナラティブを抽出する。各分類タグ（189種類）に対して、同様に推計する。

- 経済ナラティブの指数化／可視化を行う。

- ⑥推計結果をもとに、経済ナラティブインデックスとして指数化する（集計方法は後述）。
- ⑦定量評価に基づき、重要なトピック間の経済ナラティブを選択し、ネットワーク関連図として可視化する。

経済ナラティブの抽出（原因因果と結果因果の結合方法）

- （原因因果）環境規制 → （結果因果）企業戦略の経済ナラティブの場合

環境規制（原因因果）

原因のテキスト ⇒ 結果のテキスト

先行する情報
（過去情報）

現在の企業戦略の情報とつながりのある過去の環境規制の情報を抽出するため、
「原因因果の結果のテキスト」と「結果因果の原因のテキスト」が似ているペアを結合する。

企業戦略（結果因果）

原因のテキスト ⇒ 結果のテキスト

後発する情報
（現在情報）

結合したペアを「環境規制から企業戦略への経済ナラティブ」と呼ぶ。

環境規制から企業戦略への経済ナラティブのテキスト例

■ (原因因果) 環境規制 → (結果因果) 企業戦略の経済ナラティブ

環境規制(原因因果)

2006年8月

環境規制の原因:

排出権の発行を最終的に承認する国連が判断基準の引き上げ、排出権量の削減、再審査などに持ち込む例が相次いでいる

環境規制の結果:

日本企業が海外で手掛ける温暖化ガス排出権の取得事業に想定外のリスクが浮上している。

2007年11月

環境規制の原因:

日本では経済産業省や日本経団連が自主行動計画にこだわり、削減義務を伴う排出権取引を拒んできた。

環境規制の結果:

国内排出権市場の制度設計が放置され、削減の取り組みが十分に広がらないままになっている。

企業戦略(結果因果)

2008年7月

環境規制の結果

= 企業戦略の原因:

企業に排出枠を設ける排出量取引の導入議論が高まるなかで、もう一段の対策が必要

企業戦略の結果:

新日鉄と神戸製鋼は環境提携へ踏み込んだ。

経済ナラティブの指数化

■ 気候変動の原因因果と結果因果を抽出し、経済ナラティブインデックスを下記のように作成。

➤ **原因因果** (\vec{i}_{t-d}) <t時点からd日前の過去情報> とつながる **結果因果** (\vec{j}_t) <t時点の情報> の因果チェーンのコサイン類似度 (テキスト情報がどれだけ近い内容か) を計算し、月別に合計する。

ー \vec{i}_{t-d} 、 \vec{j}_t は、BERTを用いて、**原因因果**・**結果因果**のテキスト情報をベクトル表現したもの。

➤ 最新ニュースほど過去とのリンクが増えるバイアスを除去するため、経過時間に従い古い因果チェーンが減価させる (ロジスティック関数に基づき5年で重みを半減)。

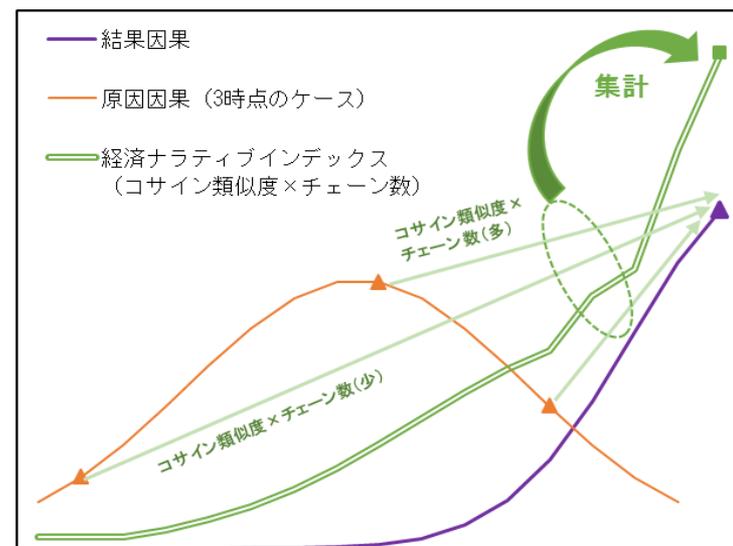
$$Index_monthly_m = \sum_{i=0}^M \sum_{j=0}^{L(i)} \frac{1}{1 + ae^{bd}} \cos(\vec{i}_{t-d} \cdot \vec{j}_t)$$

$$\cos(\vec{i}_{t-d} \cdot \vec{j}_t) = \frac{\vec{i}_{t-d} \cdot \vec{j}_t}{|\vec{i}_{t-d}| |\vec{j}_t|} \quad \text{ただし、} \cos(\vec{i}_{t-d} \cdot \vec{j}_t) \geq 0.7 \text{ を}$$

「強い因果」とみなし集計する。

(参考) t時点の結果因果につながる原因因果の因果チェーンの集計イメージ

(チェーン数)



M: m月に含まれる因果チェーンの集合

$L(i)$: 原因因果 i と接続している結果因果 j (因果チェーン)の集合

$\cos(\vec{i}_{t-d} \cdot \vec{j}_t)$: 原因因果 \vec{i}_{t-d} と結果因果 \vec{j}_t のBERTに基づくコサイン類似度

t-d: 結果因果につながる原因因果の観測時点 ($d > 0$)

t: m月に含まれる結果因果の観測時点

d: 原因因果と結果因果の時点差 (日数)

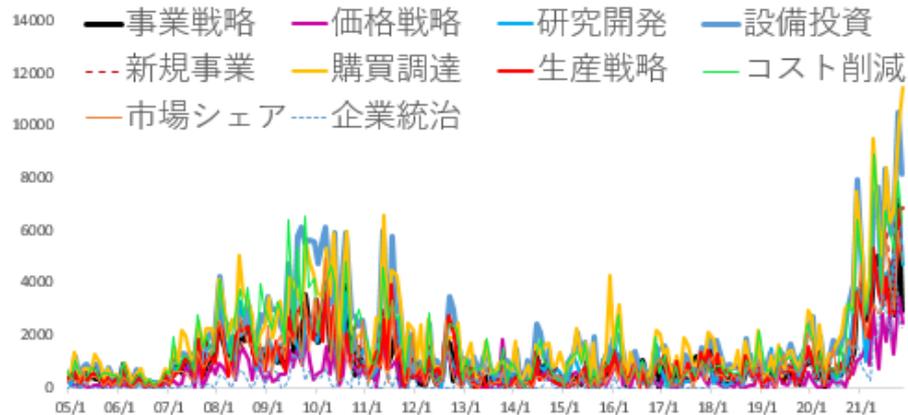
(注) ロジスティック関数のパラメータ a , b はニュースの減衰期間に応じて設定。

➤ 全月次で同様に計算し、経済ナラティブインデックスを集計 (データ期間は2000/1月~2021/11月)。
「因果チェーン数が増加する、コサイン類似度が上昇する」 ⇒ **経済ナラティブインデックスは上昇する**

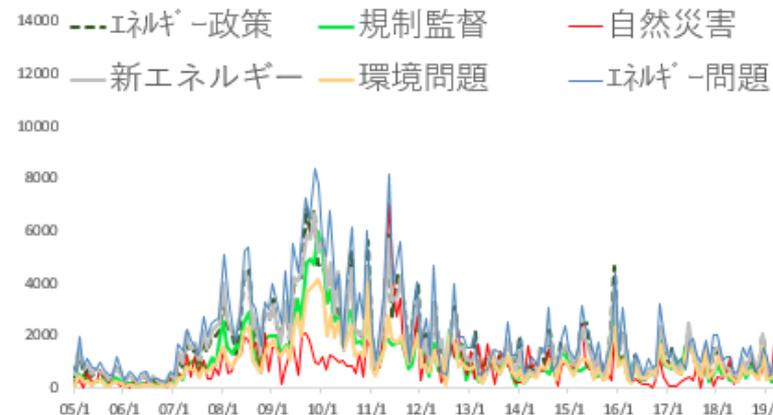
重要な経済ナラティブの探索 (40トピック)

■ (気候変動関連の) ②⑥首脳会議 ⇒ 各トピックへの関連度合にはテーマ別にみると濃淡がある。

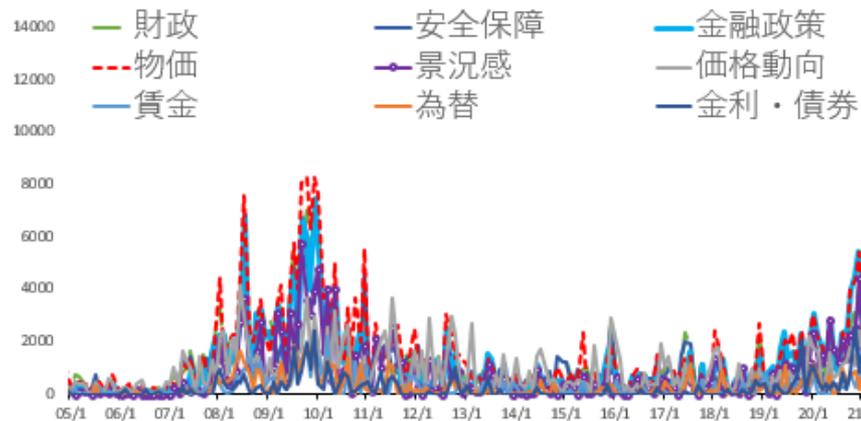
企業関連



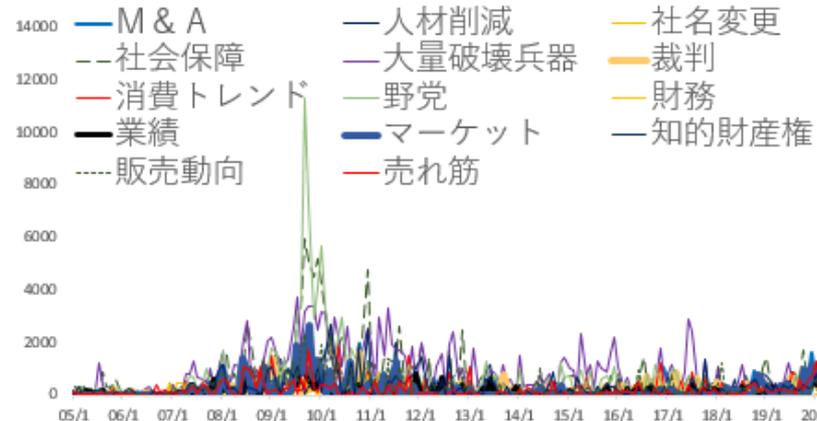
規制・制度関連



マクロ指標・政策

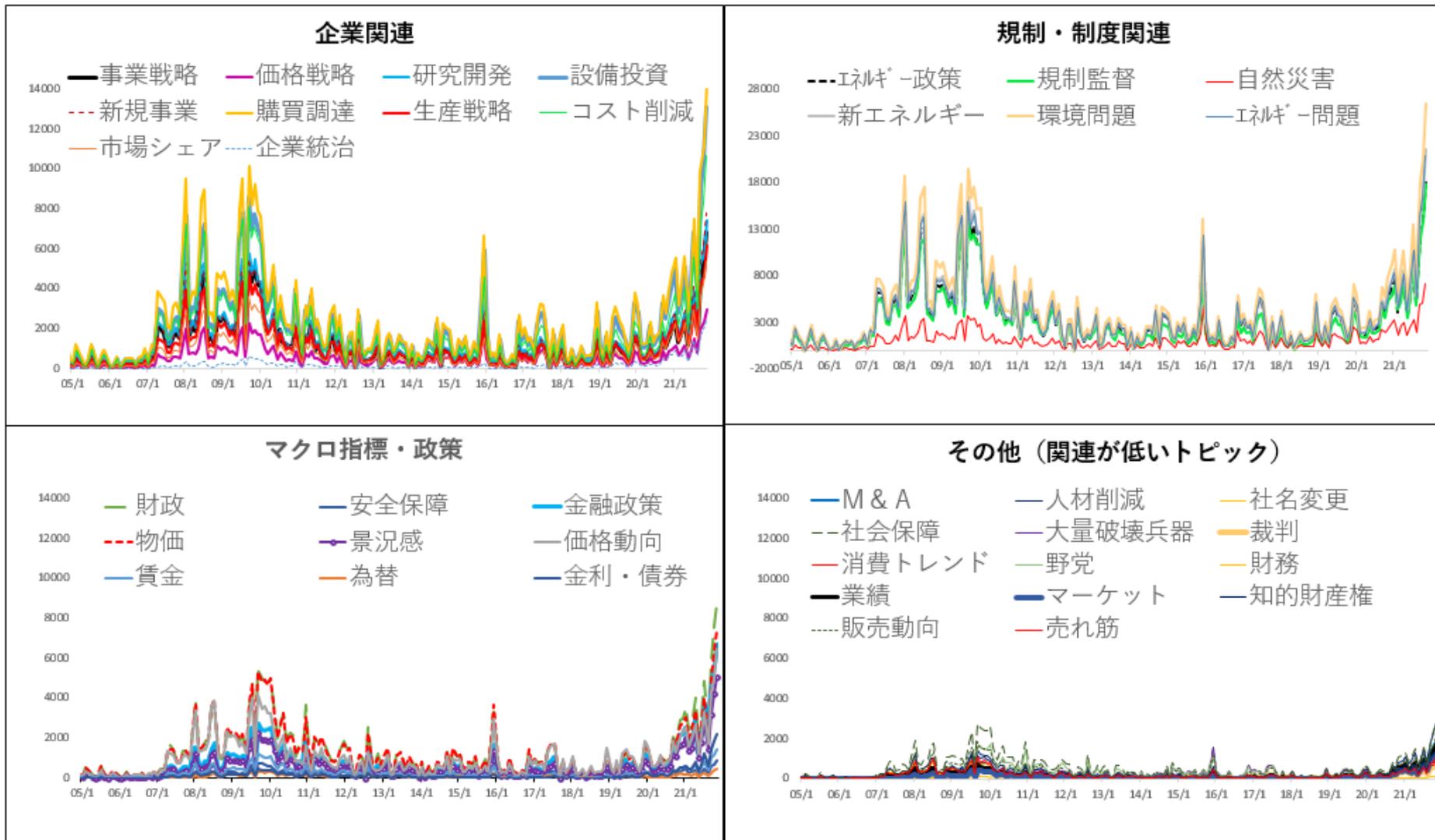


その他 (関連が低いトピック)



重要な経済ナラティブの探索 (40トピック)

■ 各トピック ⇒ ②⑥首脳会議への関連度合(逆方向)も概ね同様の傾向。ただし、規制・制度関連から②⑥首脳会議への関連は正方向より強い(国内から国際議論へのフィードバックが生じているといえるか)。



気候変動の経済ナラティブのトピック間のつながり

【主要なトピック間の因果方向】

(1) 国際的気候変動政策の進捗と環境政策・エネルギー政策強化、国内政策から国際議論へのフィードバック

原因：国際会議（グローバル）⇒ **結果**：国内環境規制・エネルギー政策（マクロ）

原因：国内環境規制・エネルギー政策（マクロ）⇒ **結果**：国際会議（グローバル）〈フィードバック〉

・ 2018年以降に顕著になった動き

(2) 国内の環境政策・エネルギー政策の強化を受けたミクロの企業行動の変化・深化

原因：国内環境規制・エネルギー政策（マクロ）⇒ **結果**：企業の事業戦略、設備投資、サプライチェーン（ミクロ）

— ただし、**企業収益や業績との連関は相対的に小さなつながり**

原因：企業の事業戦略、設備投資、サプライチェーン（ミクロ）→ **結果**：国内環境規制・エネルギー政策（マクロ）

〈フィードバック〉

(3) 自然災害と国際的な気候変動政策・環境政策・エネルギー政策の連関が出現

原因：自然災害 ⇒ **結果**：国際会議（グローバル）、国内環境規制・エネルギー政策（マクロ）

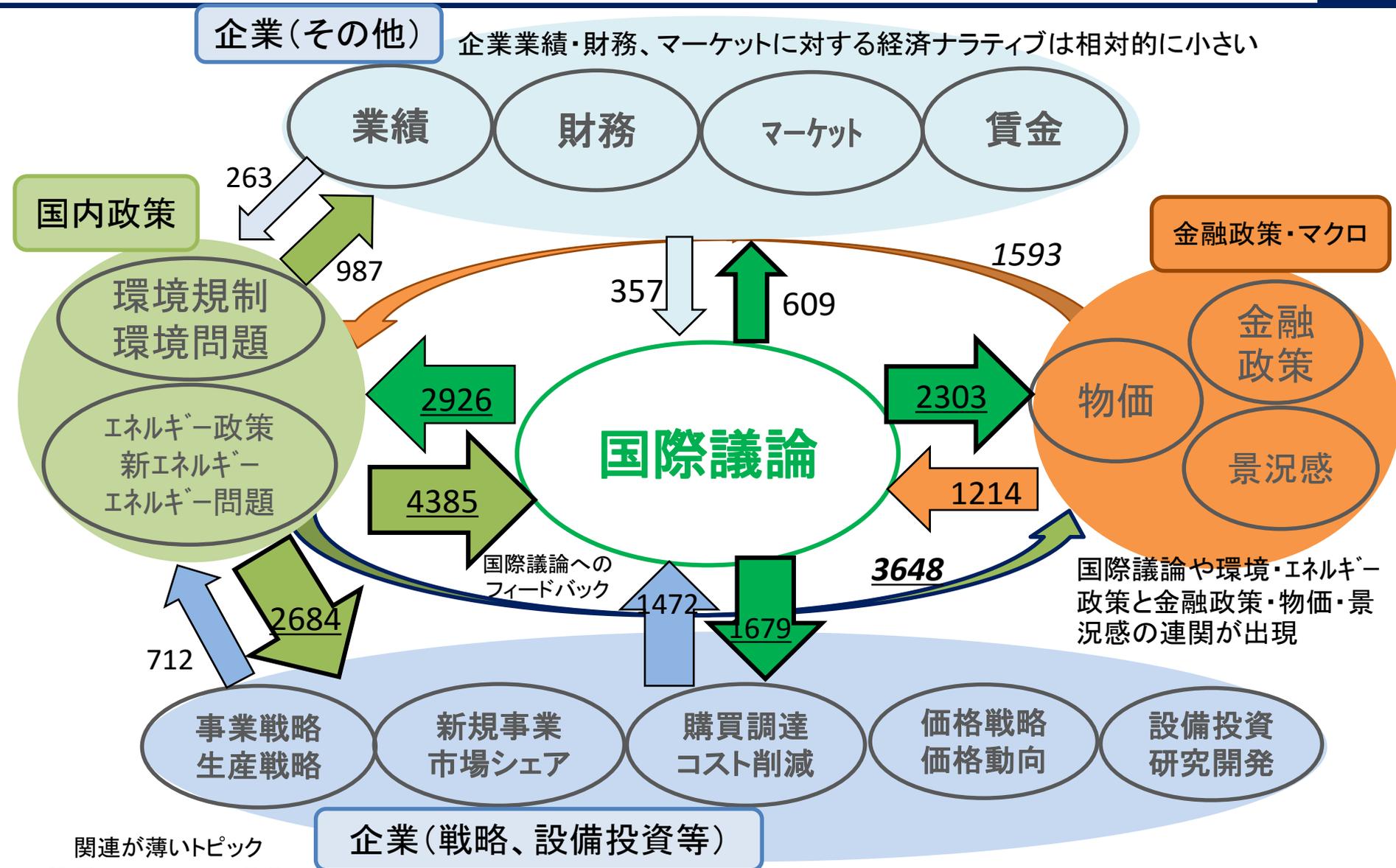
原因：国際会議（グローバル）、国内環境規制・エネルギー政策（マクロ）→ **結果**：自然災害

(4) 国内外の気候変動政策と金融政策・物価・景況感の連関が出現

原因：国際会議（グローバル）国内環境規制・エネルギー政策（マクロ）⇒ **結果**：金融政策、物価、景況感

原因：金融政策、物価、景況感 → **結果**：国際会議（グローバル）、国内環境規制・エネルギー政策（マクロ）

経済ナラティブの可視化（国際議論） <2018-21年平均>

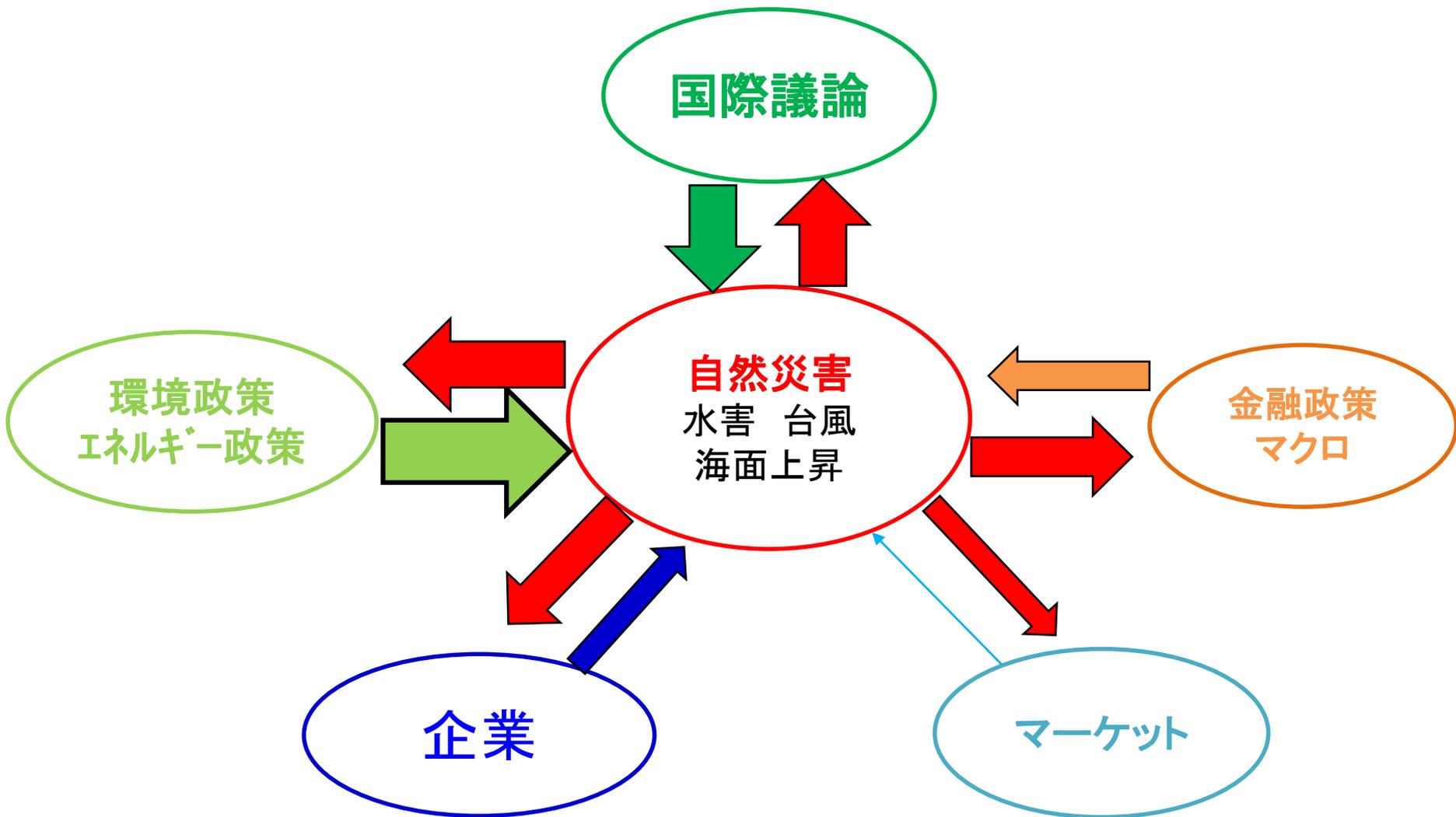


国際議論と比べ、環境・エネルギー政策から事業戦略、サプライチェーン、設備投資に対する経済ナラティブは大きい

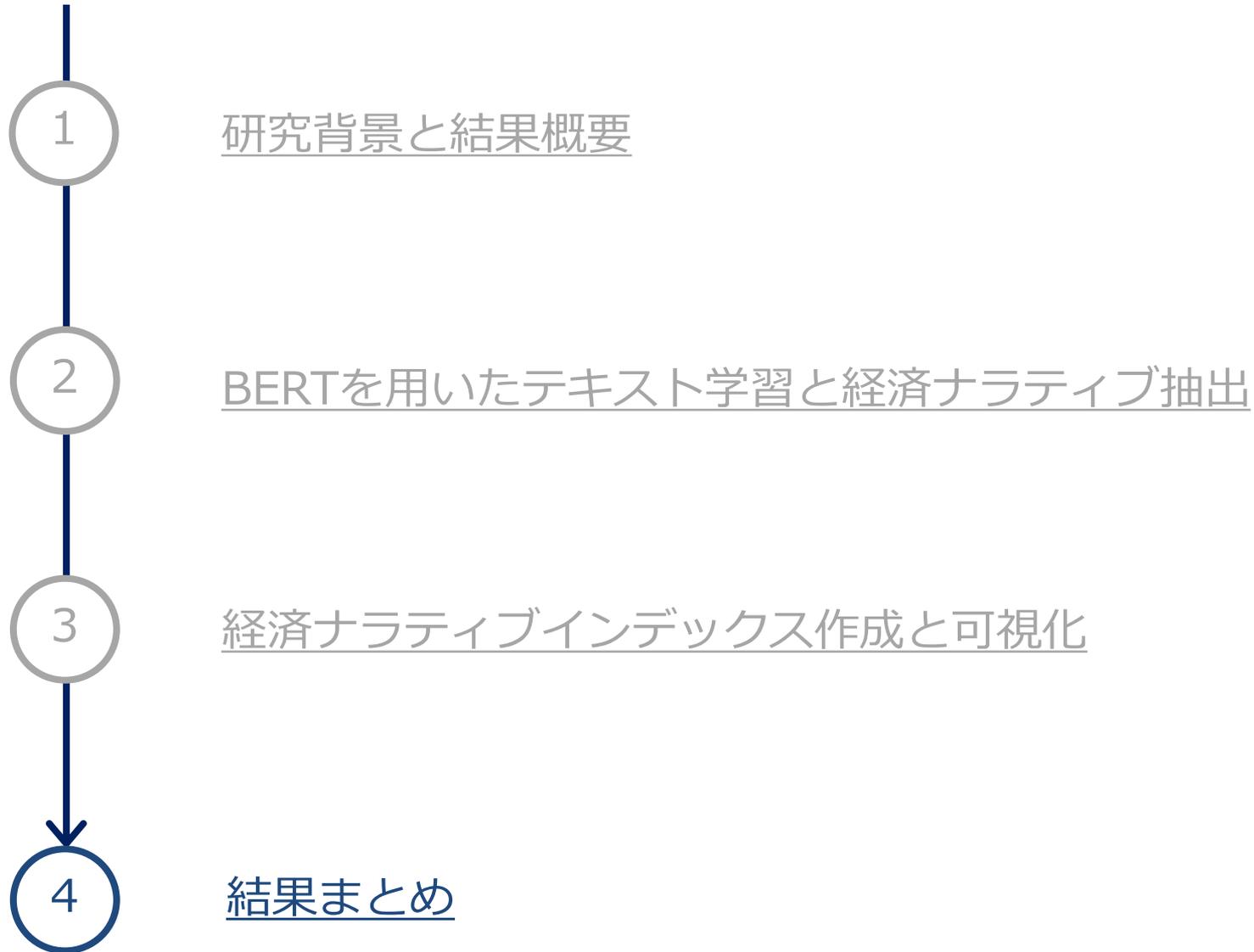
(注) 矢印内の数値は、国内政策、企業、金融政策・マクロの各カテゴリーに属する各トピックの経済ナラティブインデックスを集約し、算出した平均値。

経済ナラティブの可視化（自然災害） <2018-21年平均>

- 自然災害から各トピックへの経済ナラティブは近年増加。



(注) 矢印の大きさは、国際議論、企業関連などのカテゴリーに属する各トピックの経済ナラティブインデックスを集約し、算出した平均値の水準に相当。



気候変動の経済ナラティブの解釈と今後の課題

- 気候変動の経済ナラティブの強まりは、これまで大局的・抽象的な政策議論に止まっていた気候変動対応が、環境規制・制度として具体化されるに従い、ミクロ経済主体の認識変化や行動変化、経済・物価への影響、あるいは政策対応の検討にまで広がっている可能性を示唆。
- ✓ 移行リスク関連では、国際会議や国内環境規制の進捗や具体化する中、近年、企業や金融機関などのミクロ経済主体への経済ナラティブが強まっており、企業等の行動変化を示唆している。
- ✓ 物理的リスク関連では、自然災害の激甚化や頻度増加に関連する経済ナラティブが強まってきており、規制当局や企業経営は新たな論点・課題として認識し始めている。
- ✓ 気候変動と金融政策・物価・景況感の間の連関が近年出現している。気候変動が経済・物価に与える影響に関する議論や調査研究の増加、金融政策による対応検討を反映している可能性。

(今後の課題)

- 気候変動の経済ナラティブ（テキストにおけるつながり）が統計的因果を有するか（グレンジャー因果、構造的因果など）の検証。
- 経済ナラティブインデックスを用いた、市場データの実証分析（例えば、Engle et al. [2020] は、テキスト分析で気候変動インデックスを作成し、グリーン株のリターンとの関係を実証）。
- 今回は「気候変動ナラティブ」を分析したが、本研究は金融経済の広範なトピックを分析可能である汎用性の高さが特長であるため、今後、他の経済ファイナンス研究への応用を検討。
 - 例えば、「未来表現（～だろう）」を手掛かりに、経済主体の予想に関する経済ナラティブを抽出する手法に拡張し、インフレ予想や景気見通しに関する研究に活用できる可能性。

ご清聴ありがとうございました。

照会先

東京大学大学院工学系研究科

坂地 泰紀 sakaji@sys.t.u-tokyo.ac.jp

日本銀行金融研究所

金田 規靖 noriyasu.kaneda@boj.or.jp

参考文献

- 和泉潔・坂地泰紀・佐野仁美、2021年、「因果情報を用いた経済数値予測」、人工知能学会研究会資料、SIG-FIN-027
- 和泉潔・坂地泰紀・松島祐康、2021年、「金融・経済分析のためのテキストマイニング」、岩波書店
- 岡谷貴之、2022年、「深層学習 改訂第2版（機械学習プロフェッショナルシリーズ）」、講談社
- 岡野原大輔、2022年、「ディープラーニングを支える技術：「正解」を導くメカニズム」、技術論評社
- 翁邦夫、2022年、「「期待への働きかけ」とナラティブ：異次元緩和による事例研究」、大妻女子大学紀要、29巻、pp 69-79
- 斎藤康毅、2018年、「ゼロから作るDeep Learning②：自然言語処理編」、オライリー・ジャパン
- 坂地泰紀・酒井浩之・増山繁、2015年、「決算短信PDF からの原因・結果表現の抽出」、電子情報通信学会論文誌 D, J98-D(5), pp.811-822.
- 鈴木雅弘・坂地泰紀・平野正徳・和泉潔、2021年、「金融文書を用いた事前学習言語モデルの構築と検証」、人工知能学会研究会資料、SIG-FIN-027
- 土橋諒太・中田和秀、2021年、「BERTを用いた有価証券報告書からのESG関連文抽出」、人工知能学会研究会資料、SIG-FIN-026
- 坪井祐太・海野裕也・鈴木潤、2017年、「深層学習による自然言語処理（機械学習プロフェッショナルシリーズ）」、講談社
- 前原義明・久々宇篤志・長部喜幸、2021年、「特許ドメイン特化型BERT による脱炭素関連特許技術の「見える化」」、人工知能学会研究会資料、SIG-FIN-027
- 毛利研・春日剛・成宮仁・大場久永・関本和、2022年、「脱炭素トピック毎における相対的な進捗度合いの可視化」、人工知能学会研究会資料、SIG-FIN-028
- 若本亮佑・鶴飼和渡・里中裕輔・高木幸雄・ファイサル・ハディプトラ、2022年、「BERT を用いたESG 関連文章のGRI スタンドアード分類」、人工知能学会研究会資料、SIG-FIN-028

- Baker, S. R., N. Bloom, and S. J. Davis. (2016). “Measuring economic policy uncertainty”, *Quarterly Journal of Economics* vol131, pp.1593-636.
- Bingler, Julia Anna, Mathias Kraus, Markus Leippold and Nicolas Webersinke. (2022). “Cheap Talk and Cherry-Picking: What ClimateBert has to say on Corporate Climate Risk Disclosures,” *Finance Research Letters*, Volume 47, Part B.
- Devlin Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. (2018). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” arXiv:1810.04805.
- Engle Robert F, Giglio Stefano, Kelly Bryan, Lee Heebum, and Stroebel Johannes. (2020). “Hedging Climate Change News”. *Review of Financial Studies*, vol. 10, pages 125-152.
- Izumi Kiyoshi and Sakaji Hiroki. (2019). “Economic Causal-Chain Search using Text Mining Technology,” In *Proceedings of the First Workshop on Financial Technology and Natural Language Processing*, pages 61–65.
- Jung, Engle and Berner. (2021). “Climate Stress Testing”. Staff Reports, No. 977, Federal Reserve Bank of New York.
- Kölbel, J.F., Leippold, M., Rillaerts, J. and Wang, Q. (2021). “Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure,” Swiss Finance Institute, Research Paper Series No21-19, Swiss Finance Institute.
- Mittal Shruti, Anubhav Chauhan and C. K. Nagpal. (2022). “Stock Market Prediction by Incorporating News Sentiments Using Bert,” *Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough*, pp 35–45, Springer
- Robert J. Shiller, (2019): “Narrative Economics: How Stories Go Viral and Drive Major Economic Events”, Princeton University Press
- Hiroki Sakaji, Satoshi Sekine and Shigeru Masuyama. (2008). “Extracting Causal Knowledge Using Clue Phrases and Syntactic Patterns,” 7th International Conference on Practical Aspects of Knowledge Management, pp.111-122.
- Sonkiya Priyank, Vikas Bajpai and Anukriti Bansal. (2021). “Stock price prediction using BERT and GAN,” arXiv:2107.09055.
- Vaswani Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. (2017). “Attention Is All You Need,” arXiv:1706.03762.
- Tomas Mikolov, Kai Chen, Greg Corrado and Jeffrey Dean. (2013). “Efficient estimation of word representations in vector space,” arXiv:1301.3781.
- Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. (2018). “Deep Contextualized Word Representations,” In *NAACL2018*.