

IFC-Bank of Italy Workshop on "Data science in central banking"

18-20 February 2025

Indexing and visualization of climate narratives using BERT and causal extraction¹

Noriyasu Kaneda,
Bank of Japan

Hiroki Sakaji,
Hokkaido University

¹ This contribution was prepared for the workshop. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.



BANK OF JAPAN

Indexing and Visualization of Climate Narratives Using BERT and Causal Extraction

4th IFC and Bank of Italy Workshop
on “Data Science in Central Banking”

2025/2/19

Bank of Japan

The Hokkaido university

Noriyasu Kaneda

Hiroki Sakaji

*The opinions expressed in this document are those of the authors alone and do not represent the official views of the Bank of Japan. Any possible errors belong to the authors personally.

- “Climate change is a *global challenge* and could have a broad impact on our society and economic activity into the future.”

- BOJ’s Strategy on Climate Change

- Economists in central banks are focusing on how climate change will *affect financial markets and macro-economy*.
- Some researches used traditional text analysis to *capture climate-related risk* from newspaper and creating climate index using the NLP methods.

Engle et al. [2020], Pastor et al. [2022], Faccini et al. [2023], Hiraki et al. [2025]

- We proposed a method to analyze cause-effect relationship in text data (we call it “economic narrative”), using two NLP techniques, (1) BERT and (2) causal extraction.

(1) BERT (Bidirectional Encoder Representations from Transformers) is a Deep Learning based language model by Google, and it can learn the sentence structure and context in a text.

(2) We also use another NLP, “Causal Extraction” (Sakaji et al. [2008]). This method is an algorithm based on linguistic knowledge.

- This research focuses on climate narratives. We analyze the newspapers on climate addressing by using the NLP methods above, and index/visualize climate narratives.

【 Nikkei newspaper (morning edition, weekdays only) 】

➤ Period: January 2000 - November 2021

➤ 17,000 **climate change-related articles*** in Japanese

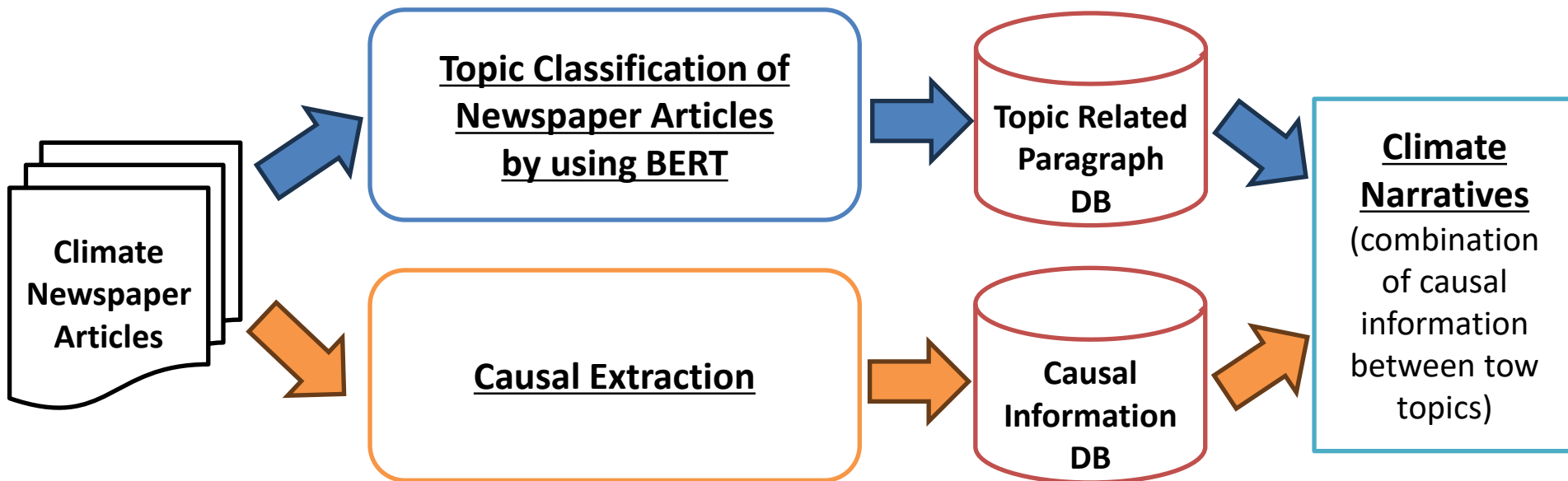
*containing at least one of the words "climate change," "global warming," or "greenhouse effect."

■ 40 topics (classification tags) are the target of this analysis.

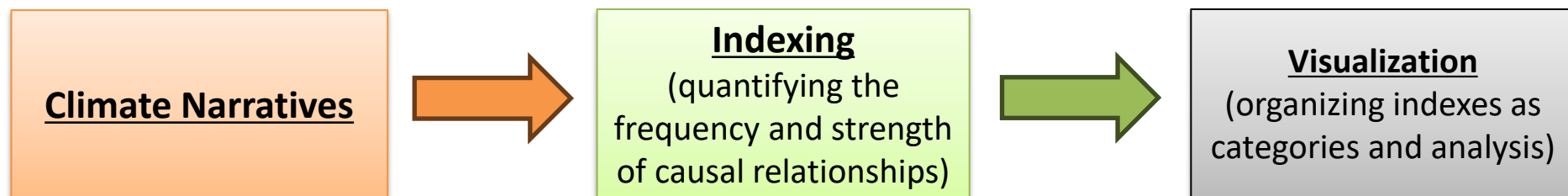
【Corporate】		【Politics】	【Economics】	【Society】
1 New business	11 Governance	21 Party	28 Monetary policy	35 Weapon
2 M&A	12 Labour	22 Regulation	29 Inflation	36 Disaster
3 Business strategy	13 Name change	23 Fiscal policy	30 Business cycle	37 Trial
4 Price strategy	14 Wage	24 Energy policy	31 Financial market	38 Energy problem
5 Production strategy	15 Finance	25 Security	32 Foreign Exchange	39 Environmental problem
6 Cost reduction	16 Performance	26 Summit	33 Bond	40 Consumer trend
7 Supply chain	17 Sales	27 Social security	34 Renewable energy	
8 Patent	18 Price			
9 R&D	19 Market share			
10 Investment	20 Hot selling			

Methodology for Extracting Climate Narratives

Methodology for Extracting Climate Narratives

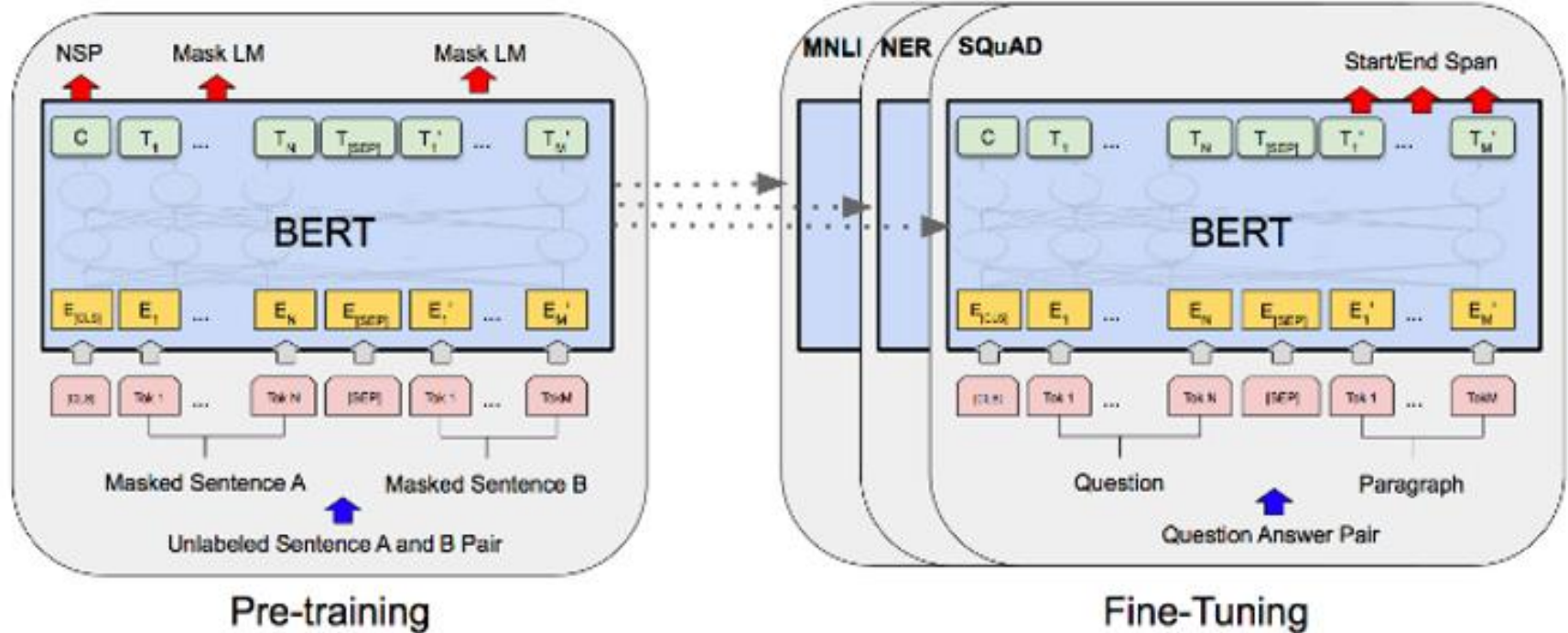


Methodology for Indexing and Visualization of Climate Narratives (for empirical research)



BERT (Bidirectional Encoder Representations from Transformers)

Devlin et al. [2019]



Trained and fine-tuned by Google
for treating basic linguistic problems
(public source)

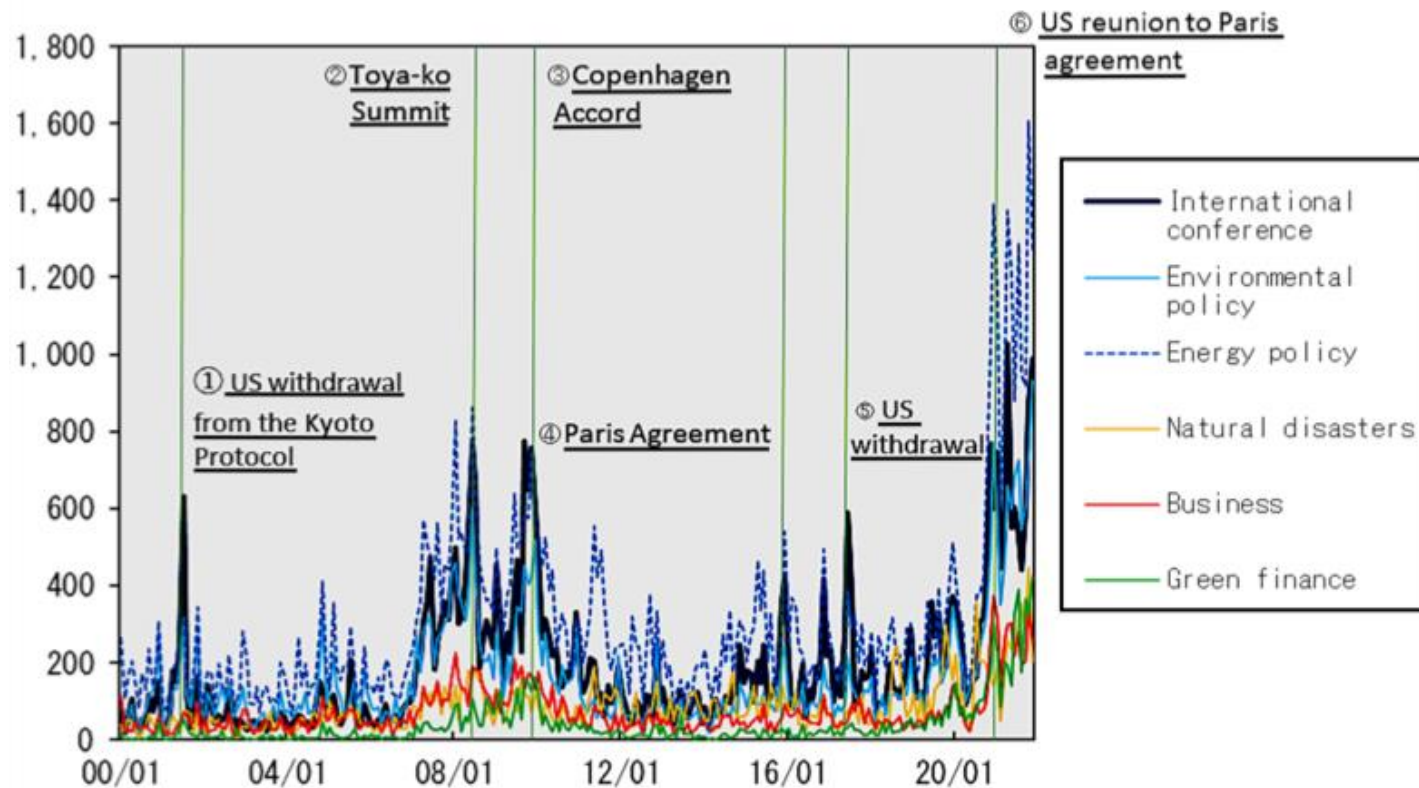
**More Fine-tuned for
classification of climate
news by topics
(specific purpose)**

	Accuracy rate	Precision rate	Recall rate	F1 score
Linear Regression	0.958	0.958	0.958	0.958
Random Forest	0.950	0.951	0.950	0.950
Support Vector Machine	0.954	0.954	0.954	0.954
BERT	0.960	0.960	0.960	0.960

- In the classification for Olympic article as an example, BERT is slightly better than other models.

Basic analysis (Topic based index using only BERT)

- **Topic based index** increased through 2008 to 2010 (in Toya-ko Summit, Copenhagen Accord) and in 2015(the Paris Agreement).
- In recent years (since 2018), **many topics** show a surge, including micro developments such as corporate strategy.



⇒ High correlation among topics. It is not clear how climate narratives are linked across topics.

- We focus on linguistic causal-effect relationship and consider the causal relationships in newspaper as economic narratives by using causal extraction.

Step.1 Causality Determination [Sakaji et al. 2017]

- Determine whether or not the input sentence contains a causal relationship (a causal sentence).

Step.2 Causality Extraction [Sakaji et al. 2008]

- Extract “cause and effect expressions” from causal sentences that contain causal relations picked up in Step1.
- This method analyzes syntactic patterns based on dependency parsing in linguistic knowledge.

Example of Climate Narrative (from Regulation to Corporate topic)

Causal sentence 1 on Environmental Regulation in Jan 2016

With Paris agreement agreed to in 2015, climate change policies are being discussed around the world.

As a result, the Japanese government plans to strengthen the environment to promote decarbonization policies.

High cosine similarity (similar information)

↓ Causal sentence 2 on Corporate Strategy in Feb 2017

In response to the government's move to tighten regulations, corporates are begging to incorporate climate change addressing into their corporate strategies.

We observe a climate narrative from Regulation to Corporate in Feb 2017.
If there are many narratives in Feb 2017, this narrative is a hot topic at this point.

Causal Event (Topic A)

Causal expression $\Rightarrow$ ***Effect expression***

Prior causal information
(past)

High cosine similarity

Combine a pair of "Causal event's *effect expression*" and "*effect expression's causal expression*": *A result relates Topic B is caused by some event relates Topic A.*

Effect Event (Topic B)

Causal expression $\Rightarrow$ Effect expression

Subsequent
causal information
(present)

Indexing Economic Narratives

- Links between **Cause** ($\vec{i}_{t-d}$) <past information from t time to d days ago> and **Effect** ($\vec{j}_t$) <information as of time t> are calculated based on the cosine similarity (how close the textual information is in content) of the causal chain and sum it by month.

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

$$\cos(\vec{i}_{t-d} \cdot \vec{j}_t) = \frac{\vec{i}_{t-d} \cdot \vec{j}_t}{|\vec{i}_{t-d}| |\vec{j}_t|} \quad (2)$$

Here,

M : set of causal chains included in month m .

$L(j)$: set of cause event $\vec{i}_{t-d}$ connected to result event $\vec{j}_t$.

$t - d$: observation point of cause event leading to result event ($d > 0$).

t : observation point of result event included in month m .

d : time difference (days) between cause event and result event.

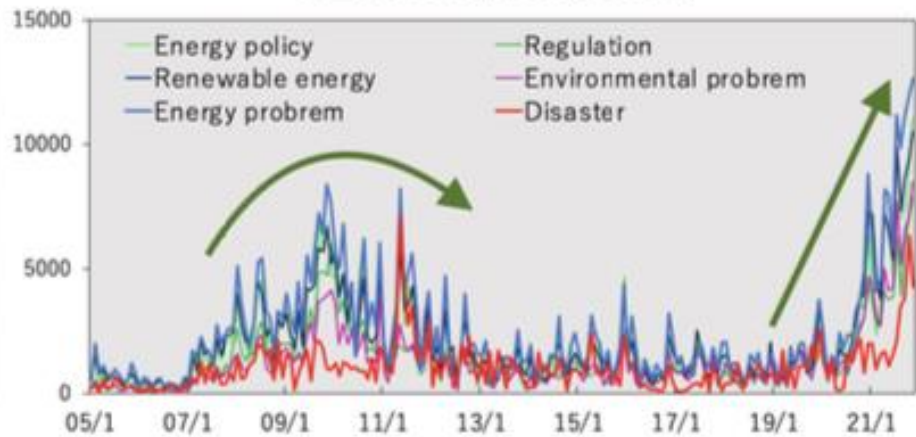
*Old causal chains are depreciated according to elapsed time to eliminate the bias that the latest news has more links to the past (weights are halved after 5 years based on the logistic function) in (1). **Count all links with $\cos(\vec{i}_{t-d} \cdot \vec{j}_t) \geq 0.7$** (strong causality) in (2).

Climate indexes (International discussion $\Rightarrow$ each topics)

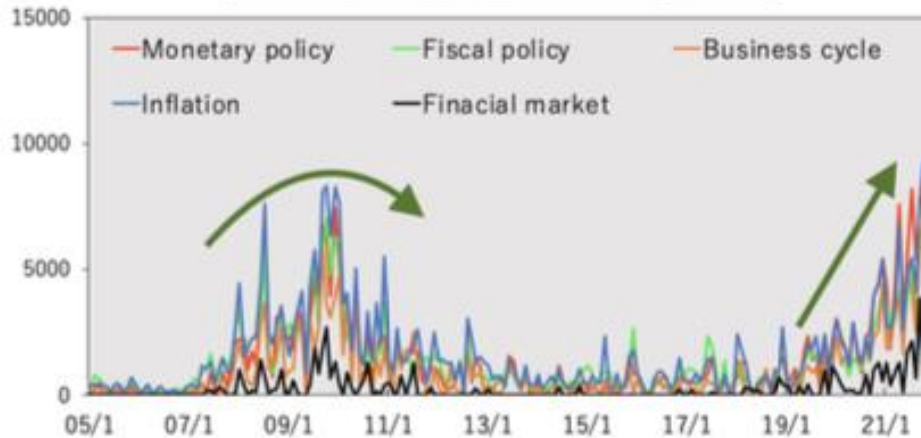
International Discussion(ID) causes Business matters



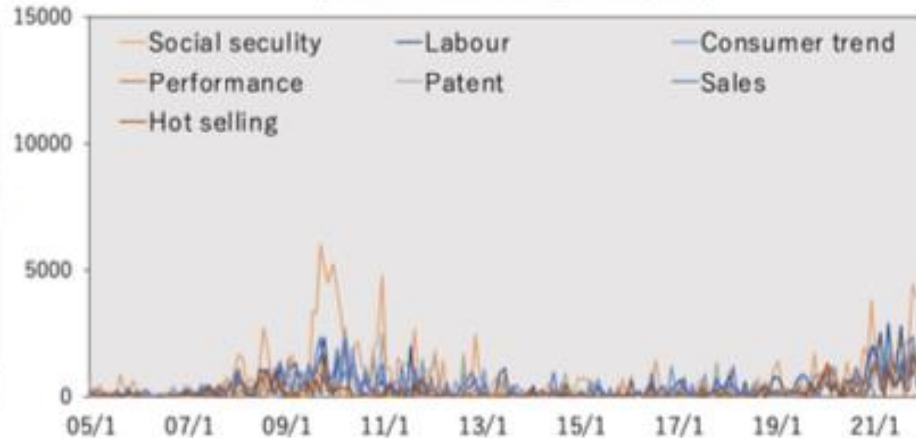
ID causes Regulatory matters



ID causes Macroeconomics/Policy matters



ID causes Other topics slightly



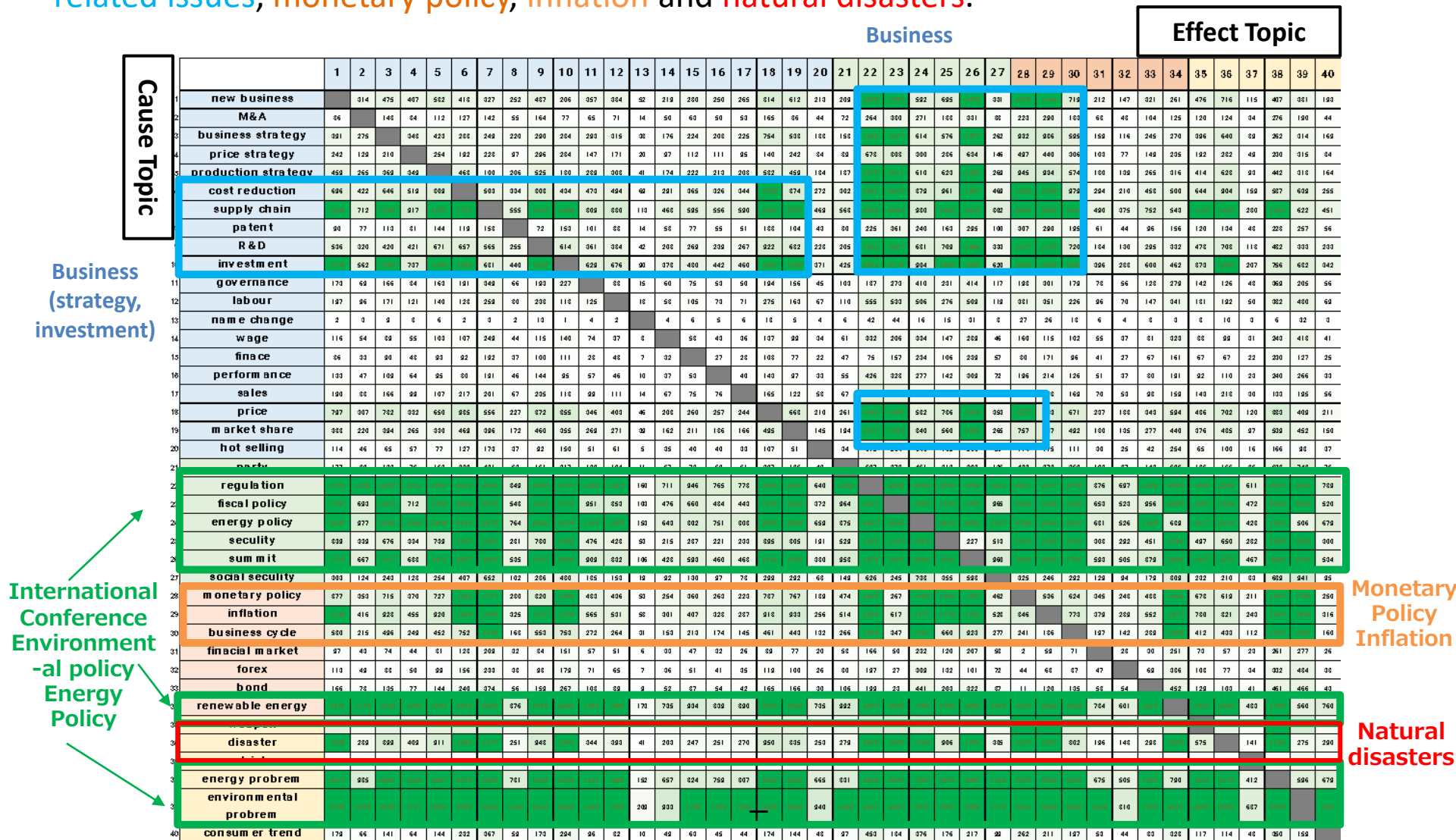
Effect Topic



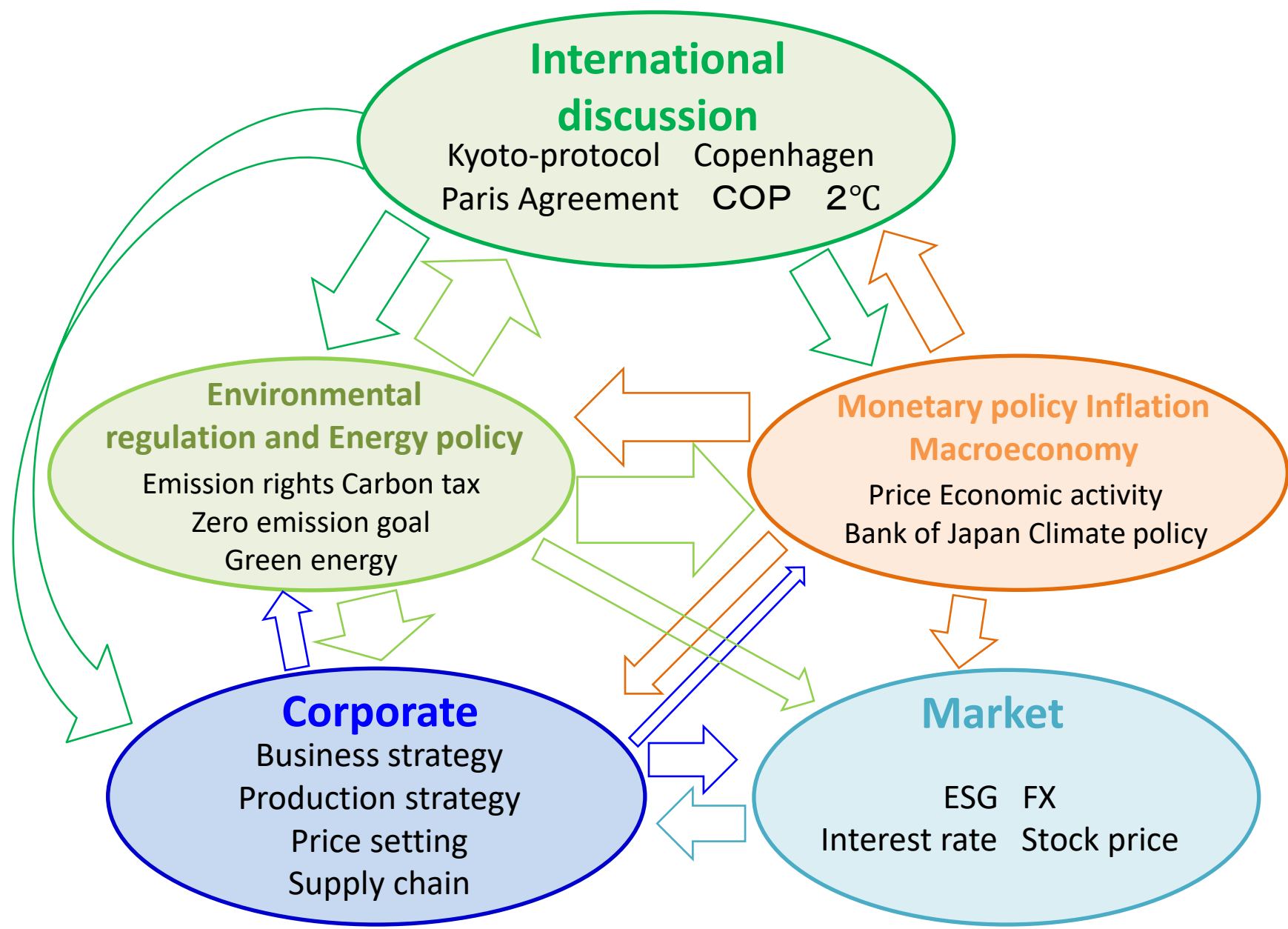
Note: Dark green highlight means a high level of economic narrative index between each topic.

Mapping all narratives from 2018 to 2021

- Strong connections on a wide range of topics from international conferences on climate change, environmental policy and energy policy. Connections have emerged in business-related issues, monetary policy, inflation and natural disasters.



Note: Dark green highlight means a high level of economic narrative index between each topic.



Note: The size of an arrow means a level of economic narrative index between each topic categories.

- The climate narratives suggest that **only governments** played a major role in the discussion of international regulations for carbon neutrality in the post-Kyoto period, or before the Paris Agreement.
- In recent years, climate change has become a hot topic not only in international and domestic regulations, but also **in corporate strategy, green finance and big issues, and as a major issue for central banks** since 2018.
- These developments could be related to the growing awareness of **transition risk** and **physical risk** in Japan.

- Is climate narrative from causal sentence statistical causality?
 - ✓ Newspaper could contain noisy or incorrect information due to the misunderstanding of the article writers.
 - ✓ Empirical analysis of market data using climate narrative indexes is the next step (like Engle et al. [2020], Faccini et al. [2023]).
 - A possible next topic: how central bank narratives about climate risk have affected the behavior of other economic entities?
- “An economic narrative is a contagious story that can change how people make economic decisions” (Shiller [2020]).*

- J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre- training of deep bidirectional transformers for language understanding,” in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pp. 4171–4186.
- R. F. Engle, S. Giglio, B. Kelly, H. Lee, and J. Stroebel, “Hedging climate change news,” *The Review of Financial Studies*, vol. 33, no. 3, pp. 1184–1216, 2020.
- R. Faccini, R. Matin, and G. Skiadopoulos, “Dissecting climate risks: Are they reflected in stock prices?” *Journal of Banking & Finance*, vol. 155, p. 106948, 2023.
- K. Hiraki, N. Kaneda, T. Kimata, and T. Mtsue, “Climate change news indices: Are they reflected in japanese stock prices?” , IMES Discussion Paper Series, 2025, forthcoming.
- R. J. Shiller, *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press, 2020.
- L. Pastor, R. F. Stambaugh, and L. A. Taylor, “Dissecting green returns,” *Journal of Financial Economics*, vol. 146, no. 2, pp. 403–424, 2022.
- H. Sakaji, S. Sekine, and S. Masuyama, “Extracting causal knowledge using clue phrases and syntactic patterns,” in *Practical Aspects of Knowledge Management*. Springer Berlin Heidelberg, 2008, pp. 111–122.
- H. Sakaji, R. Muro, H. Sakai, J. Bennett, and K. Izumi, “Discovery of rare causal knowledge from financial statement summaries,” in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017, pp. 602–608.