2024 年 9 月 19 日　全 10 頁

金融経済分析を変える自然言語処理の力①

従来のテキスト分析が与えたインパクト

定性的なテキストを定量的に測る能力などに従来から大きな魅力

経済調査部　研究員　島本 高志

［要約］

- 2022 年 11 月末に ChatGPT が登場し、大規模言語モデル（Large Language Model，以下 LLM）が注目されるようになって約 1 年 10 カ月が経った。そのような中で、金融経済分野への LLM の応用は急速に進みつつある。

- ただし、テキストデータを用いて金融市場分析や景況分析、経済学の研究を行う動き自体は、LLM が登場する前から存在した。たとえば、市場や景気の雰囲気をセンチメントスコアとして定量化する手法などは、その代表的な事例といえる。また、有価証券報告書に含まれる見通しの文章が、次期の業績予想に有効とする研究などもあった。

- 本シリーズでは改めて、大規模言語モデルを含む自然言語処理とよばれる領域が、金融経済分野での分析に対して従来もたらしてきた知見を概観するとともに、LLM がどのようなインパクトをもたらしており、今後これらの動向がどのように推移しそうなのかについて、議論していきたい。

- シリーズの前半部分にあたる本稿においては、まず、LLM より前の従来型の自然言語処理が金融経済分野に応用されてきた事例を、手法と共に具体的に紹介していく。さらに、それらを踏まえて、自然言語処理のメリットや可能性について概観することで、後半部分にあたる次回のレポートで扱う LLM の議論へとつなげていく。

はじめに

2022 年 11 月末に OpenAI 社が ChatGPT を公開し、大規模言語モデル（Large Language Model、以下 LLM）が注目されるようになって約 1 年 10 カ月が経った[1]。ChatGPT は 2023 年 1 月の月間アクティブユーザー数が 100 万人を超えるほど急速な利用拡大を見せ[2]、続く同年 3 月に発表された GPT-4 は米国統一司法試験模試で上位 10%に相当する成績を出すなど[3]、更なる性能の高さで世界を驚かせた。直近では、さらに性能が向上した GPT-4o や OpenAI o1 が出るなど、その進化は著しい。

そのような中で、金融経済分野への LLM の応用は急速に進みつつある。坂地・中川（2024）によると、日本の人工知能学会の中でも金融と情報の学際領域を扱う金融情報学研究会では、ChatGPT の公開前は LLM をテーマにした研究が皆無だった一方で、公開後には約 3 割を占めるまでに急増したという。紹介されている内容も、決算情報の要約や、ESG や SDGs に関する情報の抽出など、多岐にわたる。

ただし、金融経済分野の分析において、人間が話すような言語をコンピュータ等で処理する工学的な技術体系、いわゆる自然言語処理と呼ばれる手法（LLM もこれに含まれる）が応用され始めたのは、決して最近ではない。この自然言語処理は、LLM が登場する前からも金融経済分野の分析に応用されてきた。その内容も、景気や市場に関するセンチメントスコア（人々の心理状況を指数化したもの）を作成した研究から、中央銀行の声明文のテキストなどを定量化して市場に与えた影響を推定した分析、企業の将来見通しをテキストから定量化して将来業績に対する予測力を検証した研究まで、幅広い。

そのため、本シリーズでは改めて、自然言語処理が金融経済分野での分析に対して今までもたらしてきた知見を概観するとともに、LLM がどのようなインパクトをもたらしており、今後これらの動向がどのように推移しそうなのかを、本稿と次回のレポートにおいて議論していきたい。

まずは本稿において、LLM を用いない従来型の自然言語処理が、金融経済分野に応用されてきた事例を手法と共に紹介する。次回のレポートでは、LLM による最先端の金融経済分野における分析の事例を紹介するとともに、総論として評価や今後の展望等について議論する。

従来型の自然言語処理が活躍してきたテーマとは？

ここではまず、自然言語処理の金融経済分野への応用事例として、直近 10 年ほどの従来型の自然言語処理が応用されてきたテーマや手法を概観する（**図表 1**）。ただ、そもそも自然言語処理といっても手法は幅広く、内容も専門的となりやすい。そのため、ここでは手法に焦点を当てるよりも、むしろ分かりやすさを重視して、どのようなテーマへ使われてきたのかを切り口に

[1] 専門家の間では、後述の**補表 1** における Transformer を応用した言語モデルを全て LLM として扱うケースや、パラメータ数などを用いて LLM を定義するケースなども多いが、本稿では主に GPT-3.5 以降のモデルを LLM と呼称する。
[2] Reuters（2023）“ChatGPT sets record for fastest-growing user base - analyst note”（2023/2/3）
[3] OpenAI（2024）“GPT-4 Technical Report”, https://arxiv.org/pdf/2303.08774

紹介する。手法の詳細については、必要に応じてレポート最後の**補表１**をご覧いただきたい。

中央銀行の声明の効果を分析する

まずは、中央銀行の声明を数量化して計量的な分析へ組み込んだ事例を見てみよう。たとえば、Hansen and McMahon（2016）は、米国の金融政策を決定する連邦公開市場委員会（以下、FOMC）の声明文が、市場に与える影響について研究したものだ。これはトーン分析と呼ばれる分析法を利用し、FOMC の景気に対するスタンス（トーン）を定量的に測っている。

ただ、トーン分析自体も大きな区分で、さらに何種類もの手法の組み合わせ方がある。この論文では、まず潜在的ディリクレ配分法（Latent Dirichlet Allocation、以下 LDA）という、話題の推定などに優れたモデルを使って、FOMC が出した声明文中の景気状況に言及している文章を特定している。次に「極性辞書」と呼ばれる、ポジティブと判断される単語とネガティブと判断される単語のリストを使い、これを声明文の特定部分に適用して、ポジティブもしくはネガティブに分類される単語数を計測し、両者の合計の差を総単語数で割って、トーンを指数化している。この論文では、先行研究から景気の拡大（ポジティブ）と縮小（ネガティブ）の方向感を示す極性辞書が使われた。こうして指数化された FOMC の景気認識は、従来のマクロ計量経済学的な手法の一つである時系列分析の一変数として使われている。このように、従来の分析手法に組み込む選択肢もよく採られる。なお、この論文では、声明文の中では景気認識に関する記述よりも、将来の金利に関するフォワードガイダンスの方が市場に与える影響は大きいが、両者ともに実体経済にはそれほど影響を与えない、と結論づけている。

また、風戸ほか（2019）は、日本銀行（以下、日銀）が公表する総裁記者会見要旨のテキストのうち、質疑応答を除く冒頭部分を用いて分析を行った。具体的には、内閣府の「景気ウォッチャー調査」を学習データに使い、深層学習の手法を応用した。景気ウォッチャー調査では、全国から景気に敏感な職種の約 2,000 人を選び、現状と先行きの景気について 5 段階評価で判断したものと、その判断理由を聞いた景気判断理由集というデータセットが毎月公表されている。そのため、従来の金融経済分析における自然言語処理では頻繁に使われていた。このデータを基に学習した景気判断と文章の関係などを、日銀の総裁記者会見要旨から抜粋したテキストに当てはめてトーン分析を行い、トーン・ショックと呼ばれる情報を抽出して変数にしている。

トーン・ショックとは、統計解析によってテキストから定量的に測ったその温度感（当論文では、日銀の文章から読み取れる景気認識）であるトーンの変化率から、既に明らかになっている金融経済指標の変化率等を取り除いた残差（日銀の文章から読み取れる景気認識と、指標が示す情報の差）のことを指す。この研究ではトーン・ショックを、ファイナンスの分野で頻繁に使われる時系列分析モデルの EGARCH モデルに変数として組み込んで解析した。その結果、黒田総裁の任期期間中においては正のトーン・ショックが発生すると、つまり、日銀が景気判断を強めに下すと、円のオーバーナイト・インデックス・スワップ（OIS）という金利指標の価格変動率（ボラティリティ）を低下させ、市場参加者の金利見通しに関する不透明感を低下させたと指摘している。

図表１　従来式の自然言語処理が応用された直近 10 年間での主な研究の事例

研究	主な手法	テーマと概要	主要な示唆	国
Baker et al.(2016)	辞書法，Bag of Words（BoW）	新聞記事からの経済政策不確実性指数（EPU）というセンチメントスコアの開発	辞書によってVIXと最大約0.73の相関、変数としてマクロ経済予測の精度改善	12か国
Hansen and McMahon (2016)	LDA, BoW	トーン分析による、中央銀行のコミュニケーションの市場への影響評価	ガイダンスによるショックは現在評価の影響よりも大きい	米国
大高・菅（2018）	ナイーブベイズ分類等	景気ウォッチャー調査からの、物価センチメント指数の開発	3カ月移動平均でコアCPI前年比と、1～3カ月先行で最大約0.8ほどの相関	日本
風戸ほか（2019）	深層学習モデル	トーン分析による中央銀行の景気判断と金融市場における金融政策予想の分析	金融・経済指標の主成分を取り除いた残差であるトーン・ショックは、有意に市場へ影響	日本
加藤・五島（2021）	極性辞書	非財務情報である将来見通しのトーン分析による、将来の企業業績に対する予測力の有無	財務情報を回帰して取り除いた残差である補正済トーンは有意に予測力を持つ	日本
五島ほか（2022）	極性辞書	景気分析に特化した極性辞書の作成および、同辞書によるセンチメントスコアの評価	3種の景気指標との関係を計る手法それぞれで、景気指標と強い関係	日本
Seki et al. (2022)	BERT, サポートベクターマシン（SVM）	新聞からのサポートベクターマシン（SVM）による外れ値除去と、BERTによるセンチメントスコアの算出	景気ウォッチャー調査のDIと約0.888の相関、消費者態度指数と0.848の相関	日本

（出所）参考文献（10 ページ掲載）より大和総研作成

有価証券報告書を定量的に利用する

また、トーン分析は、中央銀行の声明の分析にしか使われないわけではない。たとえば、風戸ほか（2019）と似たような手法を扱った研究に、加藤・五島（2021）がある。彼らは、補正済トーンと呼ばれる、トーンから統計解析によって財務情報などの成分を取り除いた残差（トーンのうち、財務情報による統計モデルで説明できなかった部分）を使用することで、有価証券報告書の将来見通しのテキストを定量化し、それが将来業績の予想に有用かを検証した。

有価証券報告書のテキスト自体を企業業績予想に活用しようという動き自体は過去からあるが、たとえば単に文字数などを測るといった原始的な手法で行われることが多かった。一方で、加藤・五島（2021）は、従来から企業業績の将来予想に使われてきた定量的な指標である各種財務指標などに対して、有価証券報告書のテキストに含まれる経営者の将来見通しには、将来に関する非公開情報、たとえば企業買収や新商品開発なども加味した上での判断が掲載されていると考えた。この研究の結論では、将来に関する非公開情報が含まれている可能性の高い補正済トーンは、企業の将来業績を予測するのに有意であったとしている。

不確実性指数や物価の先行指標の作成

ここまでの研究は、FOMC の声明文や経営者の将来見通しなど、同一主体によるテキストでの継続的な発表を基に、市場や業績の変化を分析することを目的にしていた。だが、文章を定量化する手法には、他にも先行指標の構築や市場全体の心理の数量化など、多様な目的で行われたものがある。

たとえば著名な研究は、Baker et al.（2016）による経済政策不確実性（EPU）指数だ。この

研究では、時期による記事総数の差を統計的に補正しているものの、基本的には 1985 年以降の米国の代表的な新聞 10 紙で、「経済」「不確実」「赤字」「ホワイトハウス」などの単語を含む記事数を測っただけの、かなり単純な手法が使われている。だが、興味深いことに、一部の単語を金融向けに「株価」などの単語に置き換えて測った記事数で指標を作成すると、金融市場のボラティリティ予想を示す VIX 指数（時に恐怖指数とも呼ばれる）との間に約 0.73 の相関が見られたと報告されている。研究では、米国だけでなく英国や日本も含む 12 カ国に適用されて指数が作られており、中にはロシアのような報道の自由が制限されている国もあるが、その場合も EPU は市場の不確実性を測る指標として有効と確認された。論文では、政治的なスタンスが指数の測定を歪める可能性も考えて各新聞を政治的な主張に基づき右派と左派に分類し、それぞれで指数を比較しているが、両者には 0.92 の相関が見られ、連動性が高いことが報告されている。

また、前出の極性辞書を用いずにセンチメント指数を作った興味深い事例としては、大高・菅（2018）の物価センチメント指数が挙げられる。彼らは、まず、約 17 年分の景気ウォッチャー調査からランダムに抽出した 1,500 個のテキストデータを、著者の判断で「物価上昇」「物価下落」「物価もちあい」[4]「その他」の 4 つに分類した。これをナイーブベイズ分類器という手法で学習させて機械学習モデルを作成し、それを用いて景気ウォッチャー調査の景気判断理由集の全テキストデータ約 40 万個を上記の 4 種類に分類し、「物価上昇」と「物価下落」に分類されたテキストデータの数の差を物価への言及がある単語の総数で割って、センチメントスコア（人々の心理を指数化したもの）を作成した。興味深いのは、この数値に後方移動平均などの統計処理を行うと、消費者物価指数の中でも重要視されるコア指数[5]やコアコア指数[6]と呼ばれる指数との間に、数カ月間先行して最大でおよそ 0.8 の相関が見られる点だ。既存の物価統計を使わずに、テキストデータから物価の先行指標を作成した点が特徴的である。[7]

月次指標を日次指標として再現する代替指標

これまでも多くの場合に機械学習の材料として使われてきた景気ウォッチャー調査のテキストデータは月に一度の公表であるのに対し、これを日次ベースで再現しようとするような研究も存在する。たとえば、Seki et al.（2022）による S-APIR 指数は、景気ウォッチャー調査のテキストデータを学習させて、日次でのセンチメントスコアの構築を目指した。具体的には、外れ値処理にサポートベクターマシン（SVM）という手法を用い、さらに文章のスコア化に当時最先端だった BERT という自然言語処理のモデルを使用して、日本経済新聞の記事に当てはめてスコアを算出している。結果として、作成したスコア[8]と景気ウォッチャー調査の DI との間に、0.888 の相関が見られたと報告されている。また、消費マインドを示す内閣府の「消費者態度指数」との間にも、0.848 の相関が見られたことが報告されている。

---

[4] 「物価もちあい」とは、物価関連への言及はあるが、特に方向感が定まらないものを指す。
[5] 総合指数から変動の激しい生鮮食品を除いたもの。
[6] 総合指数から変動の激しい生鮮食品やエネルギーを除いたもの。
[7] なお、大高・菅（2018）は他にも、後述の共起ネットワーク図を使用して、景気変動の要因を探る手法の発展などにも貢献している。
[8] 景気ウォッチャー調査と比較するために、作成した日次指数を月次化したもの。

自然言語処理を応用した文章の定量化が金融経済分析での新領域を開拓

ここまで、個別具体的にどのようなテーマに分析が用いられてきたのかを概観してきた。自然言語処理は、たとえば金融経済分析と関係ない一般的な事例だと、文章とその書き手をペアにして機械学習にかけて両者の関係の特徴を学ばせることで、未知の文章を与えた時に文章の書き手が誰かを推測させるといった応用事例がある。あるいは、複数の文章の類似性を計算させることで、読み手に次の閲覧推奨記事を表示させるような応用事例も知られている。

では、テキストデータは、どのように金融経済分析へ応用できるとまとめられるのだろうか。

図表 2　自然言語処理が金融経済分析にもたらしたメリット

Ash and Hansen (2023)
（1）文書間の類似性の測定 （2）テキストに含まれる経済的概念の測定 （3）テキスト中の概念の関連性の測定 （4）テキストと定量的メタデータの関連付け
坂地･中川（2024）
（1）効率化 自動化により大量のテキストデータを迅速に処理し、重要な情報を素早く抽出 （2）精度の向上 経験の浅い担当者あるいは専門家と比して分析精度を向上 （3）新たな洞察 従来の人手による分析では見過ごされていたパターンを発見 （4）リスク管理 ニュース記事やソーシャルメディアなどのリアルタイムデータを適時に分析し、 市場の変動等のシグナルを早期検出 （5）カスタマイズ 顧客固有のニーズに合わせて、分析をカスタマイズ

（注）Ash and Hansen（2023）の翻訳は大和総研。

（出所）Ash and Hansen（2023）、坂地・中川（2024）より大和総研作成

金融経済分野への応用可能性を示す 4 つの「測定」

金融経済分野への自然言語処理の応用可能性に関しては、**図表 2** の上段にまとめた、Ash and Hansen（2023）における、自然言語処理を経済学の研究へ応用する際の 4 つの「測定」の定義が分かりやすいだろう。以下、解説や例などを補いながら紹介したい[9]。

まず、(1) の「文書間の類似性の測定」は、一般には先述の閲覧推奨記事の紹介などに使われる方法の応用だろう。金融経済分野の分析ならば、たとえば新規産業に関連する複数の会社が出した特許の文章を比較し、その内容の類似性を定量的に測ることで新規産業の範囲を判断し、

[9] 以下の鍵括弧内の引用部における下線による強調は、著者による

該当する企業を新興産業関連銘柄として新しい株価指数を作成することなどに向いている。このように、異なる主体の関係を定量的に利用したいときに有用だ。

次に、(2) の「テキストに含まれる経済的概念の測定」は、ネガポジ判定や感情分析とも呼ばれる、アンケート評価を判断する際など人間の心理を定量化したいときに使われる手法の応用といえる。これを金融経済分野へ応用すれば、センチメントスコアの作成のように、特定の経済事象に対する社会の関心の程度を定量的に測定することができる。

(3) の「テキスト中の概念の関連性の測定」は、共起ネットワークと呼ばれる手法が例として分かりやすいだろう。アンケートなどで任意の単語がどのくらい同時に出てきやすいかを図にすることで、テーマごとに共通する頻出単語の関連を見ることができるものだ。これを経済分析に応用するならば、新聞記事等で景気に言及した文章を解析することで、月次の景気変動の要因を探ることができるだろう。

最後に、(4) の「テキストと定量的メタデータとの関連付け」は、求人情報のテキストデータから実際の給与を予想するというような、テキストから他のデータを予測する際に役立つということである。

いずれも、自然言語処理を金融経済分野へ応用することで、従来の金融経済関連のデータでは可視化が難しかった領域を、定量化することに役立っている。つまり、テキストデータを定量的な数値データにして測れるようになること自体が、自然言語処理の大きな強みである。

実務も視野に入れた自然言語処理を用いるメリット

一方、実務に近い立場から、より広範にわたる自然言語処理のメリットをまとめたものとして、たとえば坂地・中川 (2024) がある。彼らは、金融経済分野において自然言語処理を用いるメリットとして、**図表 2 の下段の**ように、(1) 効率化、(2) 精度の向上、(3) 新たな洞察、(4) リスク管理、(5) カスタマイズ、の 5 点を挙げている。

このうちの (1) の効率化や (2) の精度の向上に関しては、たとえば、レポートを読んで内容を分類するような、従来は数時間かかる作業を機械学習で行えば、これまではコストが見合わなくて不可能だった分析も可能になり得る。また、次回のレポートで詳しく紹介する予定だが、企業アナリストのコンセンサス予想を、LLM や従来型の深層学習によるモデルの予測が精度面で上回るケースが挙げられる。これらは、自然言語処理を使うことで従来は人間が行ってきた労働を代替できる、というメリットだといえる。

次に、(3) の新たな洞察、(4) のリスク管理については、たとえば、既存のサーベイが存在しない領域に代替指標を作り出すような取り組みが考えられる。また、何らかの経済的なショックが発生した際にテキストデータを活用することで、消費マインドの急激な悪化をより早く察知できるようになることも考えられる。これらは、自然言語処理を用いることによって、従来はできなかった新しい領域が開拓できるというメリットだ。

最後に (5) のカスタマイズに関しては、たとえば、個人顧客向けに運用報告書の自動生成な

どのサービスを展開するなど、顧客向けサービスへの組み込みなどが考えられる。これは、金融経済分析へ自然言語処理を応用する場合、より実務に近い立場から見たメリットといえる。

後半にむけて

従来の自然言語処理の手法によるテーマ別の事例を見てきたが、このように、経済指標に対する先行指標や代替指標の開発、あるいは人々の経済の見方に関する心理状況の数量化や、中央銀行の声明文や企業の有価証券報告書を定量的に測ることなど、様々な目的に自然言語処理は使われてきた。現在、これらの手法の一部はLLMに取って代わられようとしている。一方、従来手法にはLLMにはない良さなどもあり、必ずしも全てが代替されるわけではないと考えられる。そのため次回のレポートでは、LLMによる最先端の応用事例等について紹介するとともに、今回紹介した従来手法と比較しながら、両者の利点や欠点を評価したのち、全体的なまとめとして、今後の展望等について議論する。

補表１　金融経済分析に応用されてきた自然言語処理の手法一覧

手法名	概要	特徴
Bag of Words	文章内に登場する単語の数をカウントする、きわめて原始的な手法。	解釈性は高いが、予測などに組み込んだ際の精度は高くなく、そのままベクトル化して使うと情報量が大きすぎる。語彙数が多い場合は計算コストが増大する傾向がある。
TF-IDF	一文内における単語の出現頻度と、単語が含まれている文の自体の希少性を掛け合わせて、単語を数量化する手法。	単語間の類似性などの算出に優れているが、精度自体は現代的にはあまり高いとはいえないケースも多い。
極性辞書	あらかじめ作っておいた関連単語のリストである辞書に対して、評価などの重みづけの点数である極性を付与しておき、文章内の登場数などと掛け合わせて文章をスコア化する手法。	景気や金融など、分野ごとに辞書を作る必要があり、辞書の精度に全体の精度が左右される。
LDA（潜在的ディリクレ配分法）	ディリクレ分布という確率分布に基づいて文章のトピックから単語が出現しているという、統計学的な考えを用いて計算するモデル。	文章を指定した数のトピックに分類することや、特定のトピックに属する文章を判別することに優れる。
Word2Vec	窓関数と呼ばれる関数を用いて、目標になる単語の周辺単語から目標となる単語の出現確率を計算する（CBOW）ことや、反対に単語から周辺の単語の出現確率を予測する（Skip-Gram）で、文章を数量化する、2013年に登場した手法。	単語の意味を周辺の単語が与えているという言語学的仮説（分布意味仮説）に基づき、周辺の単語との関係も含めて数量化することに成功した、初期のモデル。ただし、多義語や、語順の違いによる意味の差には弱い。
CNN（畳み込みニューラルネットワーク）	ニューラルネットワークの畳み込み層と結合層と呼ばれる二つの隠れ層を用いて、文脈の情報も用いて文章を数量化しようとするモデル。	画像処理などに利用される一般的なニューラルネットワークモデルであるCNNを、ニューラル言語モデルとして応用したもの。
RNN（再帰ニューラルネットワーク）・LSTM・ELMo	ニューラルネットワークのリカレント層という隠れ層で、単語同士の関係性を前後関係を考慮しながら学習することで、文脈を利用して文章を数量化するモデル。記憶セルと呼ばれる層に遠くの単語の情報を保存するとLSTM、双方向型LSTMを使用して単語埋め込みを生成するとELMoと呼ばれる。	時系列データの学習のために作られた一般的なニューラルネットワークモデルであるRNNを、ニューラル言語モデルとして応用したことで、単語の前後関係が取り込めるようになった。LSTMでは記憶セルへ情報を保存したことで、「遠くの単語の情報が劣化していく」という問題も改善した。
BERT	Attentionという、遠くの単語の情報を劣化させずに掛け合わせて保持する手法を利用する、Transformerと呼ばれる仕組みを採用した、2018年に登場したモデル。	LSTMでも完全には克服されなかった遠くの単語の情報の劣化や、前の単語の計算が終わっている必要がある問題を、並列に計算可能かつ遠くの単語の情報も劣化しなくなるなど克服し、精度が向上した。ただ、計算量の関係で短い文章にしか直接適用できず、長い文章には適用しづらい。
GPT-4（大規模言語モデルの1種）	Transformerを採用し、大量の言語データから莫大なパラメータを用い学習を行うほか、基の学習内容を超える結果を出せるように、インコンテキスト学習という明確に訓練されていないことにも能力を発揮できるようにする学習法を用いている。AIがより人間の価値観に沿った回答を出力するように、人間のフィードバックによる強化学習（RLHF）も行われており、革新的な能力を発揮しているモデル。	かなり長めの文章に対しても適用できるほか、分類タスクから要約や翻訳タスクなど幅広いタスクに適用可能で、かつチャット形式における文章から文章への返答等にも優れる。オープンソースではないため、透明性が必要な事例においては難がある場合も。

（出所）参考文献（次ページ掲載）より大和総研作成

参考文献

Ash, E. and S. Hansen (2023) “Text Algorithms in Economics,” Annual Reviews of Economics,” 15, pp.659-688

Baker, S.R., N. Bloom, and S.J. Davis (2016) “Measuring economic policy uncertainty,” The Quarterly Journal of Economics, 131(4), pp.1593-1636

Hansen, S. and M. McMahon (2016) “Shocking language: Understanding the macroeconomic Effects of Central Bank Communication,” Journal of International Economics, 99(1), pp. S114-S133.

Lee, J., N. Stevens, S.C. Han, M. Song (2024) “A Survey of Large Language Models in Finance (FinLLMs),” https://arxiv.org/pdf/2402.02315

OpenAI (2024) “GPT-4 Technical Report”, https://arxiv.org/pdf/2303.08774

Seki, K., Y. Ikuta, and Y. Matsubayashi (2022) “News-based business sentiment and its properties as an economic index,” Information Processing and Management, 59, 102795.

大高一樹・菅和聖（2018）「機械学習による景気分析―『景気ウォッチャー調査』のテキストマイニング―」，日本銀行ワーキングペーパーシリーズ，No.18-J-8，日本銀行

風戸正行・黒崎哲夫・五島圭一（2019）「日本銀行による景気判断のトーン分析」，金融研究所ディスカッションペーパーシリーズ，No.2019-J-16，日本銀行

加藤大輔・五島圭一（2021）「有価証券報告書のテキスト分析：経営者による将来見通しの開示と将来業績」，金融研究第 40 巻第 3 号，pp.45-75，日本銀行

五島圭一・新谷元嗣・高村大也（2022）「景気単語極性辞書の構築とその応用」，自然言語処理，Vol.29 No.4，pp.1233-1253，言語処理学会

坂地泰紀・中川慧（2024）「金融・経済ドメインにおける言語処理の進展」，自然言語処理，Vol.31 No.2，pp.763-768，言語処理学会

新谷元嗣（2023）「テキスト情報と機械学習を用いた景気動向分析」，内閣府経済社会総合研究所『経済分析』，第 208 号，pp.128-145

難波英嗣（2020）「テキスト間の類似度の測定」，情報の科学と技術，70(7)，pp.373-375，情報科学技術協会

松井藤五郎・汐月智也（2017）「LSTM を用いた株価変動予測」，人工知能学会全国大会論文集第 31 回，pp.1-2

ストックマーク株式会社編，近江崇宏・金田健太郎・森長誠・江間見亜利（2021）『BERT による自然言語処理入門　Transformers を使った実践プログラミング』，オーム社