

2023年11月29日

Mizuho RT *EXPRESS*

# AI-Fed ウォッチャーの誕生？

## ～ChatGPT はどこまでエコノミストに迫れるのか～

調査部プリンシパル 小野 亮

080-1069-4672 makoto.ono@mizuho-rt.co.jp

### ■ ChatGPT 公開から 1 年、進む日本企業の利活用

2023年を振り返るには少しだけ早いですが、世界のビジネス界にとっての今年の大きな話題の一つは ChatGPT だったと言っても間違いではないだろう。一般公開から 1 年。日本企業の間でも導入の動きが広がっているようだ。国内最大級の AI ポータルメディア「AIsmiley」を運営するアイスマイリーが、東証プライムに上場する 1834 社に対して行った調査によれば、9月25日時点で全体の 10% 以上である 186 社が ChatGPT 連携サービスを導入しているという。

みずほリサーチ&テクノロジーズでも、ChatGPT を利用するためのセキュアな社内環境を整備し、役職員による利活用を進めている。また API (Application Programming Interface、アプリケーション・プログラミング・インターフェイス) を社内ユーザーに開放し、ChatGPT の言わばエンジンに当たる GPT-3.5-turbo を使った、より高度・効率的なユース・ケースの開発も進められている。

API を使うことで、GPT-3.5-turbo (将来的には GPT-4) からの回答メッセージを含む大量のテキスト処理がプログラミングを通じて可能になるほか、GPT-3.5-turbo と他の情報取得の仕組み (検索エンジン等) を結び付けることで、より高度な処理も可能になる。

### ■ API 利用でエコノミスト業務も効率化

ChatGPT (あるいは GPT-3.5-Turbo や GPT-4 など) は、様々なテキスト処理・分析をこなし、業務の効率化に寄与してくれるツールである。筆者もその恩恵に与る一人だ。

そもそも ChatGPT には何ができるのか。筆者の質問に対する ChatGPT の回答を表 1 に示した。非常に多様な処理・分析をこなす可能性があることがわかるだろう。

筆者は、数年前に『退屈なことは Python にやらせよう — ノンプログラマーにもできる自動化処理プログラミング』という書籍を手にいれ、Python を実戦投入して以来、技術革新の民主化を肌で感じてきた。ChatGPT の登場はさらに劇的な変化、リープフロッグ現象をもたらしてくれている。

例えば、自分の業務効率化の一環として、YouTube の海外動画から英語の自動生成キャプションを取得し、GPT-3.5-turbo によって人間が読める (エコノミストが分析できる) 形に編集するための対話形式のシステムを構築したり、FOMC (米連邦公開市場委員会) の議事要旨について、直近と前回とを文単位で比較しながら閲覧できる仕組みを構築したりしている。

前者を開発するに至った背景は次の通りである。通常、パウエルFRB（米連邦準備制度理事会）議長を始めとした米金融政策当局者の発言が講演原稿という形でしか入手・確認できない。公開パネル討論などに参加した際の発言は、メディア報道に頼ることがほとんどであり、その結果、断片的でしかない。YouTube動画のキャプションは、すべて小文字であるため、一部をレポートに使うだけでも相当程度、編集する必要がある。また発言の文脈を理解するには、やはり動画を丁寧に確認する必要がある。

表1 ChatGPTが教えてくれたテキスト分析の例

|   |   |   |
|---|---|---|
| <ol style="list-style-type: none"> <li>1. <b>感情分析</b>: ChatGPT を使用して、テキスト・データ中の感情や意見を分析し、ポジティブなトレンドやネガティブな感情を抽出</li> <li>2. <b>キーワード抽出</b>: ChatGPT を利用して、テキスト内で頻出するキーワードやフレーズを抽出し、主要なトピックやテーマを把握</li> <li>3. <b>テキスト要約</b>: ChatGPT を活かして、大量のテキスト・データを要約し、主要な情報や要点を簡潔に整理</li> <li>4. <b>トピックモデリング</b>: ChatGPT を用いて、テキスト・データの潜在的なトピックを同定し、それに基づいて文書をクラスタリング</li> <li>5. <b>文法・言語の解析</b>: ChatGPT を使って、テキストの文法や言語構造を分析し、誤りや不自然な表現を検出</li> <li>6. <b>時系列分析</b>: ChatGPT を利用して、テキスト・データの時系列パターンを抽出し、特定のトピックやキーワードの変化を追跡</li> <li>7. <b>異常検知</b>: ChatGPT を使用して、通常の文体や内容と異なるテキストを検出し、異常なパターンを識別</li> <li>8. <b>言及ネットワークの構築</b>: ChatGPT を使って、テキスト内のエンティティやトピック間の言及関係を抽出し、ネットワークを構築</li> <li>9. <b>要素の重要度評価</b>: ChatGPT を活かして、テキスト・データ中の要素や文の重要度を評価し、重要な情報を抽出</li> <li>10. <b>意味的類似性の評価</b>: ChatGPT を使用して、テキスト・データ内の文や文章の意味的な類似性を評価し、関連性の高い要素を特定</li> </ol> | <ol style="list-style-type: none"> <li>11. <b>エンティティ抽出</b>: ChatGPTを用いて、テキスト内の具体的な個人、場所、日付などのエンティティを抽出し、情報の構造化を支援</li> <li>12. <b>言語の変化分析</b>: ChatGPTを利用して、異なる時代や地域での言語の変化を追跡し、文化的な変遷を把握</li> <li>13. <b>文脈依存の解釈</b>: ChatGPTを使って、文脈に依存した表現や隠された意味を解釈し、正確な理解を支援</li> <li>14. <b>テキストの信頼性評価</b>: ChatGPTを活かして、テキスト・データの信頼性や情報源の評価を行い、偽情報やバイアスを検出</li> <li>15. <b>言語間の翻訳支援</b>: ChatGPTを使用して、異なる言語のテキストを解釈し、翻訳の補助や文脈の適切な伝達を支援</li> <li>16. <b>オピニオンマイニング</b>: ChatGPTを利用して、テキスト・データ中の意見や評価を抽出し、製品やサービスに関するフィードバックを整理</li> <li>17. <b>メディアトピックの特定</b>: ChatGPTを用いて、メディアにおける特定のトピックの報道頻度やその変化を追跡</li> <li>18. <b>顧客フィードバックの感情分析</b>: ChatGPTを活かして、顧客のフィードバックやレビューから感情や嗜好の傾向を分析</li> <li>19. <b>特定トピックのキーパーソン抽出</b>: ChatGPTを使用して、特定のトピックにおいて中心的な発言者やキーパーソンを抽出</li> <li>20. <b>テキスト生成モデルの評価</b>: ChatGPTを活用して、他のテキスト生成モデルと比較し、性能や適用範囲に関する評価を実施</li> </ol> | <ol style="list-style-type: none"> <li>21. <b>専門用語の抽出</b>: ChatGPTを使用して、特定の専門分野における専門用語や専門的な表現を抽出し、ドメイン特有の言語を把握</li> <li>22. <b>文書の構造解析</b>: ChatGPTを活かして、テキスト内の見出し、段落、リストなどの構造を解析し、文書の階層構造を把握</li> <li>23. <b>テキスト生成の多様性評価</b>: ChatGPTを用いて生成されたテキストの多様性を評価し、モデルの表現の幅や柔軟性を検討</li> <li>24. <b>知識グラフ構築</b>: ChatGPTを活かして、テキスト・データから関連性のある情報を抽出し、知識グラフを構築</li> <li>25. <b>機械学習モデルの説明性解析</b>: ChatGPTを使用して、複雑な機械学習モデルの出力や予測を解釈し、モデルの説明性を向上</li> <li>26. <b>アノテーションの自動生成</b>: ChatGPTを活用して、テキスト・データに対する適切なアノテーションやタグを提案</li> <li>27. <b>テキストのバイアス分析</b>: ChatGPTを用いて、テキスト・データに潜むバイアスや偏りを検出し、公平性の観点から評価</li> <li>28. <b>学習データの偏りの評価</b>: ChatGPTを使用して、モデルの学習データにおける偏りや不均衡を評価し、モデルのロバスト性を確認</li> <li>29. <b>異なるドメインへの適応性評価</b>: ChatGPTを活かして、特定のドメインで学習したモデルが他のドメインにどれだけ適応できるかを評価</li> <li>30. <b>テキスト・データのプライバシー保護</b>: ChatGPTを使用して、テキスト・データから個人を特定できる情報を検出し、プライバシー保護対策を実施</li> </ol> |
|---|---|---|

(注)最初の10個は「ChatGPTを使ったテキスト分析手法を10個提案して下さい」というプロンプトへの回答。残りは「追加的に10個教えてください」というプロンプトへの回答。アウトプットはプロンプトに対応して「ですます調」となっており、上記表では筆者が末尾を編集している(出所)ChatGPTより、みずほリサーチ&テクノロジーズ作成

既存の技術とGPT-3.5-Turboを組み合わせた仕組みによって、こうした一連の作業は実務上、極めて短時間で終了する。その速度とアウトプットの精度は筆者にとっては十分なものだ。キャプションの自動取得は弊社内のネット速度次第だが1時間半の動画で10秒以下、文章の変換は1分30秒程度で済む。(そもそも、全て小文字で表記された1万数千もの英単語を、自分が読める形に手直することなど、時間があっても、自らやる人は少ないだろう。)

生成されたテキストは、話し手の入れ替わりもある程度捉えられており、当局者の発言を会話の文脈の中で理解する手助けになる。完璧なテキストにするには動画を確認する必要があるものの、キャプションと同時にタイムスタンプ(経過時間データ)を取得して処理すれば、動画確認もスムーズだ。

なお、報道されていない他の発言も含め、米金融政策スタンスの変化を捉えるためのテキスト・データとして蓄積・利用できるようになるのは言うまでもない。

議事要旨を文単位で比較閲覧する仕組みは、ChatGPTではなく、情報検索に秀でた別のAI・機械学習モデルを使っている。プログラミングによって議事要旨の取得から比較閲覧まで極めて短時間で処理できるのはそれだけで大きな成果だ。①FRBのウェブサイトから議事要旨を取得⇒②参加者の討議パートを文単位で新旧比較するエクセル・ファイルを作成⇒③当該ファイルをエクセルで開く、という一連のプロセスをプログラミングによって自動実行・終了するまで、わずか25秒程度である。

## ■ ChatGPT が示す AI-Fed ウォッチャーのポテンシャル

エコノミストの分野におけるChatGPTを使った本格的な分析として筆頭に挙げられるのは、今年の春に発表された2つの論文であろう。ChatGPTは、業務の効率化だけではなく、専門家の分析領域に足を踏み入れるだけの能力を秘めているようだ。

Bloombergは4月、次のように報じている。

「リッチモンド連銀のアン・ルンドガールト・ハンセン、ソフィア・カジニク両氏が執筆した「Can ChatGPT Decipher Fed Speak? (チャットGPTは米金融当局の言い回しを解読できるか)」と題された第1の論文によると、当局の声明がハト派のかタカ派のかの判断でチャットGPTは最も人間に近い成績をあげた。チャットGPTの解釈はグーグルの自然言語処理モデル「BERT」や辞書に基づく分類より正しかった。」

「もう一つは、フロリダ大学の研究者、アレハンドロ・ロペスリラ、ウィエファ・タン両氏がチャットGPTに金融専門家になったつもりで企業ニュースの見出しを解釈するように求めた論文「Can ChatGPT Forecast Stock Price Movement? Return Predictability and Large Language Models (チャットGPTは株価の動きを予測できるか。リターン予測と大規模言語モデル)」である。チャットGPTのトレーニングデータには含まれていなかった2021年後半以降のニュースを用いた。結果は、チャットGPTの回答がニュース以後の株価の動きと統計的に関連付けられることを示した。このテクノロジーを使ってニュースの意味合いを正しく読み取ることが可能であるという表れだ。」

(原文ママ。引用元：<https://www.bloomberg.co.jp/news/articles/2023-04-18/RT9EL1T0AFB401>)

これらの先例に倣い、筆者もGPT-3.5-TurboのAPIを使って米国の金融政策スタンスの変化をテキスト・データからどの程度読み解けるのかを検証してみた。

本稿で用いたデータは、2017年12月から2023年11月までの議事要旨である。なお米金融政策のスタンスの変化は、FOMC当日のパウエル議長による記者会見や、その後のFOMC参加者らによる講演やインタビューなどのコミュニケーションを通じて、すでにかかなりの程度市場参加者に伝わっている。FOMCから3週間後に公表される議事要旨の中にサプライズな情報が含まれることは少ない。今回の検証は、あくまで議事要旨の文章をFedウォッチャーと同じように読めるのかどうかに焦点を当てるものであることは改めて理解されたい。

FOMCの議事要旨はほぼ固定された複数のパートに分かれている。例えば2023年11月の議事要旨は参加者のリストを除き次の見出しで構成が分かれている。

- Developments in Financial Markets and Open Market Operations (金融市場動向と市場操作)
- Staff Review of the Economic Situation (スタッフによる経済レビュー)
- Staff Review of the Financial Situation (スタッフによる金融レビュー)
- Staff Economic Outlook (スタッフによる経済見通し)
- Participants' Views on Current Conditions and the Economic Outlook (FOMC参加者の討議パート)
- Committee Policy Actions (参加者のうち、政策決定の投票権を持つ“メンバー”による討議と政策決定)

本稿の分析では「FOMC参加者による討議パート」に焦点を当てた。議事要旨を読み解こうとする(人間の)エコノミストは、使われている単語の違いや、新たに追加された分析を重視し、スタンスの変化を捉える。プログラミングとプロンプトを工夫することで、GPT-3.5-Turboにも同様の作業を行わせることができるのだろうか。

## ■ 微妙な政策スタンスの変化も捕捉

AI・Fedウォッチャーの誕生は近そうだ。

分析では米金融政策のスタンスを表す7つのラベルを用意し、GPT-3.5-Turboに議事要旨の内容をラベリングさせてみた。その結果が図1である。

ここで近年の米金融政策を簡単に振り返ろう。パンデミック直前のFOMCは、予防的インフレ対策としての引き締めから、景気への配慮を重視した利下げに転じる局面にあった。2000年のパンデミック発生によって、FOMCは急激に緩和モードに切り替え、政策金利が実効下限に張り付くと、バランスシートの拡大による追加緩和を追求した。2021年、サプライチェーンが回復途上の中で需要が急回復し、それが高インフレのきっかけとなったが、FOMCがインフレへの警戒感を強めるのは同年終盤のことであった。そして実際に利上げに踏み切ったのが2022年前半であり、同年後半には大幅利上げを繰り返すに至った。2023年に入ると、インフレのピークアウトと政策金利水準がすでにかかなり高いことから、FOMCは利上げペースを緩め、7月の利上げ以降は様子見が続いている。

図1は、上記のようなFOMCの政策スタンスの変化を、GPT-3.5-TurboがFOMC参加者らの討議内容から概ね正確にとらえてラベリングしていることが分かる。

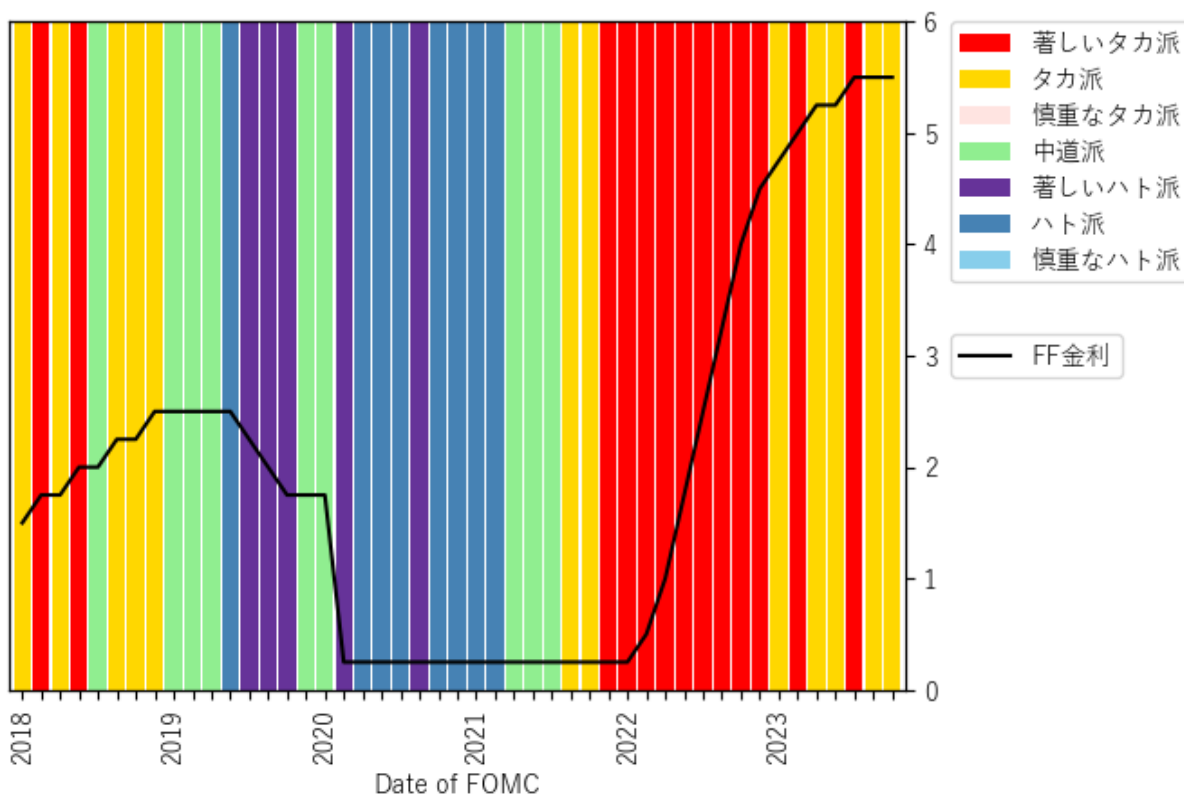
ラベリングの難しさを痛感するのは、パンデミック前の利下げ局面である。景気への配慮から利下げに転じた局面のラベリングとして筆者が期待したのは「慎重なタカ派」だが、「ハト派」を含むラベルが付与されている。景気への配慮が必要というFOMC参加者らの意見が「ハト派」として認識さ

れ、政策金利の高さに対する評価はラベリングの優先度が低いようだ。2024年にはFOMCが利下げに転じるとみられ、GPT-3.5-Turboがどのようなラベルを付けるのか、これから楽しみである。（なお実際のプロンプトは英語である。プロンプトは言わば「料理の秘密のレシピ」という性格を持つため、本文及び図1では簡単な日本語を当て直している。）

総合的にみれば、「米国の金融政策スタンスの変化をテキスト・データからどの程度読み解けるのか」という問いに対して、分析結果はポジティブな答えを示したように思われる。モデルに「AI-Fedウォッチャー」としての追加学習をさせなくても、議事要旨の内容から政策スタンスの微妙な変化を捉えることができる、という発見は少なくとも筆者には驚きであった。また追加学習させれば、「シンギュラリティ」とは言わずとも、エコノミストに相当程度近づけることは間違いないように思える。

一方、分析を進める上で課題や気づきもある。例えば、プロンプトで指定したフォーマットに即していない形で回答を得る場合だ。こうした問題が起きると、GPT-3.5-Turboの回答内容を次の分析に使っていくような、バックエンドでの自動処理を進めるシステムを作った際に、エラーを吐き出し、次の処理に対する誤ったデータを入力してしまう恐れがある。

図1 GPT-3.5-Turboによる米金融政策のスタンス分析



(出所)FRBより、みずほリサーチ&テクノロジーズ作成

気づきとしては、ユーザーが期待する通りのテキスト処理・分析を行わせるには、プロンプトで与えるタスクからも、曖昧さを極力排除した方が良いという点が挙げられる。プロンプトはシンプル・イズ・ベストというのが、筆者の実感である。本稿の分析では英語のプロンプトを使ったが（複数のプロンプトで試行錯誤した）、日本語のプロンプトを与える場合には一段の注意が必要であろう。「論理的に正しい」と思える文章でも、ユーザーの「意図」が伝わっているかどうかは別問題である。これはヒトとのコミュニケーションに通じる課題だ。

さらに、適切なプロンプトを探し当てるのはそれほど容易ではなく、エコノミストとしての知見（一般にはドメイン知識と呼ばれるもの）、プログラミング及びプロンプトエンジニアリングのスキル、そして試行錯誤を繰り返す忍耐と時間が必要である。

以上、生成AIによる分析がエコノミストの業務分野に応用できる簡単な例を示した。分析を進める上での課題は多々あるが、生成AIに対する理解や、プログラミング及びプロンプトエンジニアリングに対する理解をユーザーが深めるほど、その分だけ、得られる成果は付加価値の高いものになりそうだ。言い換えれば他社（他者）との差別化である。進歩は止まることなく、人間とAIの連携が、より効率的で高度な分析を目指す一助となることが期待される。

[本コンテンツに関するアンケートに](#)

[ご協力をお願いします](#)



**【PR】YouTube<sup>®</sup>動画「MHRT Eyes」・各種 調査レポート（無料）を配信中！**（「YouTube」は Google LLC の登録商標です）

～国内外の経済・金融動向など幅広い分野について、エコノミスト・研究員が専門的な知見をご提供～

▽メルマガ（登録無料）では、配信をいち早くお知らせしております。下記より是非お申込みください

<https://www.mizuho-rt.co.jp/publication/mailmagazine/research/index.html>



（QR コードはデンソーウェブの登録商標です）

お問い合わせ：みずほリサーチ&テクノロジーズ株式会社 調査部メールマガジン事務局

（03-6808-9022, [chousa-mag@mizuho-rt.co.jp](mailto:chousa-mag@mizuho-rt.co.jp)）

●当レポートは情報提供のみを目的として作成されたものであり、取引の勧誘を目的としたものではありません。本資料は、当社が信頼できると判断した各種データに基づき作成されておりますが、その正確性、確実性を保証するものではありません。本資料のご利用に際しては、ご自身の判断にてなされますようお願い申し上げます。また、本資料に記載された内容は予告なしに変更されることもあります。なお、当社は本情報を無償でのみ提供しております。当社からの無償の情報提供をお望みにならない場合には、配信停止を希望する旨をお知らせ願います。