

AI の発展と会計学研究

首藤昭信
(東京大学)

要 旨

本稿の目的は、AI 関連の会計学研究の主要な論文を体系的に整理し、その展開を概説することである。具体的には、(1) 機械学習を利用した財務数値の予測、(2) テキスト分析を中心とする自然言語処理の応用、そして (3) 生成 AI の出現が会計学研究にもたらす新しい潮流、という 3 つの観点からサーベイを行う。本稿では、それらの研究成果を整理することで、AI の発展が会計学の展開にどのように寄与してきたのかを明らかにする。

キーワード：AI, 機械学習, 自然言語処理, 生成 AI, 会計学

1 本論文の目的と構成

近年、AI (Artificial Intelligence) の進展は社会全体に大きな影響を与えており、会計の世界もその例外ではない。AI 技術の中核を成す機械学習 (machine learning)、深層学習 (deep learning)、および生成 AI (generative AI) は、データの収集、処理および分析のあり方を根底から変化させ、会計実務と会計学研究の両方に影響を与えている。会計研究者は AI 技術を活用することで、会計の新しい側面の解明やより精度の高い分析を行うことができるようになった。また AI の会計実務への浸透は、企業の会計情報の作成と開示行動を効率化することに加えて、投資家といった会計情報の利用者の行動も変化させている。このような AI に伴う会計環境の変容は、会計情報の意義や機能を再検討する契機を与えている。

本稿の目的は、AI 関連の会計学研究の主要な論文を体系的に整理し、その展開と学術的意義を概説することである。具体的には、AI 技術のうち、(1) 機械学習を利用した財務数値の予測に関する分析、(2) テキスト分析を中心とする自然言語処理の応用、そして (3) 生成 AI の出現が会計研究にもたらす新潮流、という 3 つの観点からサーベイを行う。本稿では、それらの蓄積を整理することで、AI の発展が会計学の展開にどのように寄与してきたのかを明らかにする。

2 予測を目的とした機械学習の利用

2.1 従来の研究からの展開

機械学習が会計学の領域に利用され始めたのは「予測」を重視する研究領域からである。例えば、企業の倒産、不正会計・修正再表示、または将来業績の予測といったテーマは、1970 年代から会計学の重要な研究領域を形成してきた。これらの研究は、投資家といった利害関係者が将来事象を予測する際に会計情報を利用しているか、すなわち会計情報の有用性を確認する重要な研究課題であった。

財務比率にもとづく倒産予測研究では、ロジスティック回帰やプロビット回帰モデルといった計量経済学の手法を用いることで、Altman (1968) の Z スコアや、Ohlson (1980) の O スコアといった予測スコアが提案されてきた。また不正会計研究では、Dechow et al. (2011) が提案した不正スコア

(F スコア) は、その後の多くの実証研究の基盤となっている。これらの研究では、研究者は会計学の理論的裏付けのある説明変数をピックアップして予測を行うことで、会計情報の意義および有用性を検討していた。これらの変数が予測事象と有意な関連性を持てば、会計情報とその背後にある会計理論の有用性を示唆する結果となるためである。

しかし、近年の企業環境は複雑化し、従来の限定的な変数によるモデルでは複雑な現象を十分に説明できないとの認識が高まってきた。機械学習の活用は、大量データと高度なアルゴリズムに依拠した分析を可能にするため、上記の会計学の予測モデルを改善する可能性を有する。例えば、Perols (2011) は、会計不正を予測する際の機械学習を中心とする 6 つのアルゴリズムの比較を行った結果、機械学習による推計が従来の手法をアウトパフォームすることを示している。また Bao et al. (2020) はアンサンブル学習 (RUSBoost) を用いて財務諸表の生データから不正を検出するモデルを構築した。その結果、従来の比率ベースの変数と伝統的なロジスティック回帰モデルを利用した分析よりも、機械学習を利用した分析は高い予測能力を示すことを示している。さらに近年の業績予想に関する研究は、機械学習の利用は業績予想の正確性を改善することも示されている (Jones et al., 2023; Easton et al., 2024)。これらの一連の研究は、機械学習を予測モデルに利用することで予測モデルの精度が向上することを示唆している。

2.2 説明可能な AI

機械学習は予測モデルの精度を向上させる一方で、結果の解釈が困難であるという問題が指摘されていた。いわゆる推計モデルの「ブラック・ボックス」の問題である。この問題に対処する 1 つの方法として注目されているのが「説明可能な AI (Explainable AI: XAI)」である。この方法は、複雑なアルゴリズムの内部構造を可視化し、意思決定の根拠を提示することを目的としており、代表的な技術として LIME (Local Interpretable Model-agnostic Explanations) や SHAP (SHapley Additive exPlanations) がある。これらの手法の利用は、変数ごとに重要度 (importance) を計算することで、それぞれ説明変数の分類能力や予測に対する寄与を定量的に示すことができる。

機械学習を用いた会計学研究の近年の最も大きな進展の 1 つは、この XAI を導入したことである。予測モデルに利用された説明変数 (特徴量) の相対的な重要度が把握できれば、会計学研究に与える含意は非常にリッチになるためである。Bertomeu et al. (2021) は、①会計、②資本市場、③コーポレート・ガバナンス、④監査、および⑤企業特性といった 5 つの観点から 100 以上の説明変数を用いて、修正再表示の発生を機械学習 (メインは Gradient Boosted Regression Tree) を利用して見抜くことを試みた。その結果、上位の変数には、従来の研究で重視されてきた会計ベースの変数だけでなく、資本市場や監査といった変数が選択されていることを示した。

Zhang Parker et al. (2025) は、将来発生する修正再表示を見抜く予測モデルを構築している。これまでの研究は、修正再表示の発生後の検出に焦点を当てていたが、本研究は将来発生する修正再表示を予想することで、企業・監査人・規制当局・投資家が事前にリスクを把握し、予防的な対応が取れるようにしている。分析を行った結果、1 年後または 2 年後に発生する修正再表示を見抜く高い検出力を有するモデルを識別するだけでなく、SHAP 分析を行うことで包括利益、外国企業か否か、未認識の税制優遇措置の未払利息とペナルティといった項目が主要な予測要因となっていることを明らかにしている。

2.3 新しい潮流

近年の機械学習を利用した研究は、従来の倒産、不正会計・修正再表示、または将来業績の予測にとどまらず、様々な研究領域の拡大を見せている。第1に、会計学の伝統的な研究領域である価値関連性研究に機械学習を応用し、会計情報の有用性を再検証する動きがある。価値関連性研究に機械学習を利用することは、会計項目と株価の関連性を非線形の関係性を考慮して捉えるだけでなく、XAIの利用を通じて大量の変数の重要度を議論できるというメリットがある。例えば Barth et al. (2023) は、CART (Classification And Regression Tree) を利用して分析した結果、利益や簿価といった伝統的な会計項目の説明力は近年低下傾向にあるものの、無形資産や成長性に関する会計項目の説明力は増加しており、むしろ会計情報全体の説明力は増加している、という興味深い結果を提示している。また Starica and Marton (2025) は、利益と株価の関係に関する複数の理論モデルについて、そのモデルが現実即しているかを機械学習を用いて識別しようとする意欲的な研究である。具体的には、利益と株価の関係についてあらかじめ関数形を仮定せず、機械学習 (Random Forest) を用いて推定を行い、理論モデルの予測形状と比較するという分析を行っている。その結果、実証研究の多くが仮定する線形モデルではなく、ダイナミック・オプションモデルの予測と最も整合的であることが示されている。

第2に、予測の精度の向上だけでなく、予測の対象が拡大し、その分析内容が深化している。例えば、近年の研究は、金融機関の貸倒期待損失引当金 (Lu and Nikolaev, 2022)、将来の実効税率 (Guenther et al., 2023)、アナリスト予想利益 (Drake et al., 2024) といった項目にも分析対象が拡大している。アナリスト予想利益の分析は多数の先行研究が存在するが、Drake et al. (2024) はアナリストの利益予想のスタイルを、機械学習を用いて5パターンに分類するという興味深い試みを行っている。例えば、クオンツ型 (ファンダメンタルなどの数量分析を重視)、コントラリアン型 (市場動向と逆の情報を重視)、またはハーダー型 (他のアナリスト動向を重視) といった分類である。分析を行った結果、予測スタイルの多様性が高いアナリストがフォローする企業では、コンセンサスの分散は大きくなる一方で、予測の精度は向上することを示している。さらにそのような予測の多様性は、企業の情報環境を改善することも例証している。このような分析は、機械学習がアナリスト予想研究に新しい知見をもたらしているといえる。

3 テキスト分析

3.1 従来の研究からの展開

会計学研究におけるテキスト分析 (textual analysis) とは、企業の開示文書や報道、アナリスト・レポートといった非構造化言語データを定量的に分析し、経営者の意図や投資家の認識を推定する研究領域である。分析対象として最も一般的なのは米国のアニュアル・レポートであり、とくに経営者による討議と分析 (Management Discussion and Analysis: MD&A) などが主な題材とされてきた。

この分野の研究は、まず記述情報の情報内容を分析する「トーン」や「可読性」の分析から始まった。トーンとは、記述内容の傾向を意味し、例えば将来業績に対して肯定的、否定的、または中立的か、といった視点から分類される。初期の研究である Bryan (1997) は、約250社のMD&Aの将来業績に関する予測 (トーン) を手作業で分類し、そのような予測は短期において実際に実現することを示した。また Li (2008) は、アニュアル・レポートの可読性を Fog Index を用いて計量化し、企業業績が悪い企業のアニュアル・レポートほど読みにくいことを示している。一連の研究は、アニュアル・レポートの定性的情報が投資家の意思決定に影響を及ぼすことを実証的に示している。

初期のトーンに関する研究は、その識別にあたり、単語の発生頻度に応じて分類を行う辞書ベース・アプローチが取られていた。Li (2010) は、このトーンの分類に機械学習を利用した初めての研究である。MD&A を分析対象として単純ベイズのアルゴリズムにもとづく分類を行い、MD&A のトーンは将来業績と関連を有することを示した。また彼は、従来の辞書ベースのアプローチでは同様の結果が得られないことも示している。Li (2010) と先行研究との最大の相違点の 1 つは、当該研究が 130 億の記述文を分類していることである。高度な分析方法に加えて、サンプル・サイズの大きさは比較にならないほど充実している。

3.2 テキスト情報を用いた予測

機械学習を利用するテキスト分析の研究は、財務数値を対象とした分析と同様に、不正会計や信用リスクといった予測を行う研究へと展開した。アニュアル・レポート内で財務諸表と一緒に開示される記述情報が、財務諸表数値と同様の予測能力を有しているかが関心となったのである。

まず、不正会計に関する研究として Purda and Skillicorn (2015) は、米国企業のアニュアルレポート（四半期報告書含む）の MD&A を対象に、サポートベクターマシンを用いて不正企業を識別するモデルを構築した。その結果、財務比率を用いた従来モデル (Dechow et al., 2011) よりも高い予測精度を達成した。ただし、記述情報にもとづく彼らの予測モデルは、財務数値にもとづく予測モデルとは異なる側面を捕捉することも明らかにしている。この結果は 2 つの予測モデルを併用することの合理性を示唆する。

また Brown et al. (2020) は、トピックモデル (Latent Dirichlet Allocation: LDA) を用いてアニュアル・レポートのトピックを抽出し、それらが不正会計（ミス・リポーティング）の発生確率を予測できるかを検証した。分析の結果、不正企業は好業績を強調する一方で、コストやリスク要因への言及が少ない傾向を示すことが確認された。

信用リスクの予測に関しては、Donovan et al. (2021) がある。彼らは、企業の MD&A やカンファレンス・コールを対象に複数の機械学習の結果を統合して生成した「信用リスク・スコア」が、伝統的な信用リスク指標（Z スコアや O スコアなど）を上回る予測力を有することを示した。格付情報はすべての企業で入手可能ではなく、彼らが推計したスコアはすべての企業に適用可能であるという点から実務的にも有用であると主張されている。

3.3 新しい潮流

近年のテキスト分析研究は、単なるトーン分析や予測分析を超え、会計情報の新しい活用方法を検討する段階へと展開している。その新しい潮流は、(1) 財務情報とテキスト情報の融合、(2) 企業特性や企業行動の新たな側面の抽出、という 2 つの観点から整理できる。

第 1 に、財務分析とテキスト分析の融合である。Kim and Nikolaev (2024) は、MD&A における記述情報と財務数値の相互作用に着目し、両者の相互作用が企業の将来パフォーマンス予測を改善するかを調査した。具体的には、深層学習 (deep learning) アプローチを用いて、①会計数値のみ、②記述情報のみ、③相互作用を考慮せず両者を使用、④相互作用を考慮して両者を併用するという 4 パターンのモデルを比較した。分析を行った結果、記述情報と財務数値の相互作用が、企業の将来パフォーマンス予測（業績・キャッシュフロー・株式リターン）の精度を高めることを明らかにした。特に財務数値の信頼性が低い場合や経済の不確実性の高い時期に、相互作用を考慮することで予測精度の

改善が顕著になることを示している。

第2に、テキスト分析を通じて企業特性や企業行動の新しい側面を抽出する研究が増加している。例えば、企業の投資機会集合 (Basu et al., 2022) やセグメント分類 (Song, 2021) の識別がテキスト分析で行えることが示されている。より最近の研究としては、監査領域でもテキスト分析の応用が進んでいる。Hope et al. (2025) は、国際監査基準 (ISA) へのコンバージェンスが監査の質に与える影響を国際比較した研究であり、概して、コンバージェンスが監査の質を高めることを示している。注目すべきは、国際監査基準と各国の監査基準とのコンバージェンスの程度について、機械学習に依拠して測定していることである。自然言語処理の技術を用いて、ISA と各国の監査基準の類似性を様々な観点から検出することでコンバージェンスの程度を測定している。

また Blankespoor et al. (2023) は、IPO ロードショー (IPO 実施前に行う機関投資家向けの説明会) における経営者プレゼンテーションのテキストを分析し、経営者が投資家に対してどのような情報を強調するか (pitch) を分析している。その結果、経営者は IPO ロードショーでは、ポジティブな表現を多用し、不確実的な表現が少ないことを示している。

さらに、ESG 分野でもテキスト分析の応用が拡大している。Lin et al. (2024) は、世界各国のアンニュアル・レポートを対象に、単語埋め込みモデルを用いて ESG 関連語彙の出現頻度を分析し、ESG の開示傾向を捉える独自の辞書を作成した。この辞書をベースにして、近年の ESG 開示は、開示量は増加しているが、その質的内容は低下していることなどを示している。

4 生成 AI

4.1 研究の体系

生成 AI (Generative Artificial Intelligence) は、テキスト、画像、音声、コードなどの新たなコンテンツを生成する AI 技術であり、社会のあらゆる分野で急速に浸透している。特に ChatGPT に代表される大規模言語モデル (Large Language Models: LLMs) に依拠したアプリケーションは、言語理解、要約または推論といった人間固有の知的活動を高精度に模倣し、その影響は教育、法務、医療のみならず、会計・監査分野にも及んでいる。実際に、規制やサービス障害等で ChatGPT が使用できない場合、アナリストの活動または株式市場の取引量などが低下することが近年の実証研究で示されている (Bertomeu et al., 2025; Cheng et al., 2025)。生成 AI が市場参加者の不可欠のツールとなっていることが示唆される。

生成 AI を扱う会計学研究の初期段階では、AI が人間の会計専門職を代替できるか、という実践的関心から出発した。たとえば、Wood et al. (2023) は GPT を会計・監査の問題に適用し、基本的な知識問題には一定の精度で回答できるものの、数値処理や監査判断においては誤答が多いことを報告している。Cheng et al. (2024) は、新しいバージョンの GPT を利用することで、既存のフレームワークを用いた倫理的判断を伴うタスクでは生成 AI のパフォーマンスは高くなることを示している。より体系的な分析を行ったのが Eulerich et al. (2024) である。彼らは、生成 AI が米国の主要な会計資格試験 (CPA, CMA, CIA, EA) に合格できるかを調査し、GPT-3.5 ではいずれの試験にも合格できなかったが、GPT-4 に Few-shot 学習や ReAct プロンプトを導入することで、全ての試験に合格する水準に達することを報告している。この結果は、生成 AI が人間の専門的知識をすでに有していることを示すものであり、会計・監査業界を脅かす結果であると主張している。

こうした初期研究の後、近年では、生成 AI を財務分析や投資意思決定といった具体的な文脈に応

用する研究が登場している。特に、アニュアル・レポートやカンファレンス・コールなどのテキスト情報を題材に、生成 AI の能力を検証する研究が進展しており、その機能は大きく「予測」「分類」「要約」「抽出・定量化」の4つに整理できる。以下では、それぞれの代表的研究を概観する。

4.2 主要な研究成果

4.2.1 予測

生成 AI を活用して財務数値や将来業績を予測する研究は、テキスト分析の流れを継承している。Li et al. (2025) は、ランダムに選ばれた 500 社の四半期決算プレスリリースを GPT-4 に読み込ませ、業績予想を行わせた。その結果、GPT-4 は「早い位置にある文章」「可読性の高い文章」「センチメントが明確な記述」などを重視して判断し、情報環境が不透明な場合には数値情報を重視する傾向を示した。また、GPT-4 による予想はアナリスト予想と比較して精度が低い、情報開示の質が高い企業では予想精度が顕著に向上することが確認された。

Jha et al. (2024) は、カンファレンス・コールのテキストを GPT-3.5 に入力し、将来の資本的支出額を予測した。その結果、GPT が生成する指標は Tobin's q などの伝統的指標をコントロールしても長期的な投資行動を予測できることが示された。さらに Kim et al. (2024) は、財務諸表の数値のみを GPT-4 に与え、Chain-of-Thought プロンプトを通じて翌期業績の増減益を予想させた。その結果、単純なプロンプトでは他の機械学習モデルを下回るが、Chain-of-Thought プロンプトを用いた場合にはアナリストやニューラルネットワークに匹敵する精度を示した。この知見は、生成 AI の数値情報の予測能力がプロンプト設計によって大きく変化することを示唆している。

4.2.2 分類

de Kok (2025) は、カンファレンス・コールの質疑応答を題材に、経営者の「非回答 (non-answer)」を GPT が識別できるかを検証した。結果として、GPT による分類は既存の機械学習モデルよりも精度が高く、生成 AI が経営者の口述内容の文脈に対する理解が高いことを示した。

また Breitung and Müller (2025) は、79,000 社超の事業内容に関する説明文を生成 AI でベクトル化し、コサイン類似度に基づく企業分類指標を構築した。この指標により企業分類を行った結果、類似企業間のリード・ラグ効果を利用した取引戦略で有意な超過リターンを得られることを報告している。さらに、Cao et al. (2025) は、生成 AI にピア企業を特定させた結果、人間の専門家が特定したピア企業や、ベンチマーク手法によるピア企業と高い一致率を示すことが確認された。また、生成 AI が特定したピア企業は、株式リターンや売上高成長率、粗利率において、対象となる企業との相関が高いことも確認された。以上の結果から生成 AI がピア企業を合理的に特定できると結論づけている。

Bilinski (2024) は、FTSE100 企業のアニュアル・レポートを GPT-4 で分析し、センチメント・スコアが株価反応や翌期 ROA と正の関係を有することを示した。同様に、Lopez-Lira and Tang (2024) はニュースヘッドラインを GPT で分類し、そのセンチメントが株式リターンを予測する上で他のモデルを上回ることを確認した。これらの結果は、生成 AI の分類能力が投資家行動分析に新たな有用性を持つことを示している。

4.2.3 要約

生成 AI の他の特長の 1 つは、長文要約能力である。Kim et al. (2025a) は、MD&A およびカンファレンス・コールを GPT で要約させ、原文と要約文のセンチメントを比較した。要約文は原文の約

20%の長さであるにもかかわらず、開示日付近の株価反応をより強く説明することが示された。また、原文と要約文の差を「膨張度 (bloat)」と定義し、文章が冗長な企業ほど価格効率性が低いことを明らかにした。また Kim et al. (2025b) は、政治リスクや気候リスクなど特定テーマに焦点を当てた「ターゲット要約」を作成し、生成 AI を通じたリスク指標が株式ボラティリティを説明することを示している。これらの研究は、要約が投資家の情報処理コストを軽減することを示唆している。

4.2.4 抽出・定量化

生成 AI を用いて非構造化テキストから指標を抽出・定量化する研究も増えている。Choi and Kim (2024) は、Form 10-K の記述から企業が税務調査を受けているかを GPT-4 で判定する指標を開発し、実際の税務当局データと高い一致率を示した。この指標を利用して、税務調査中の企業は租税回避を抑制し、ボラティリティが上昇し、投資や資金調達が縮小することを明らかにしている。

また、Bernard et al. (2025) は、GPT を利用して注記情報の XBRL タグから事業の複雑性を測定し、複雑性が高い企業では報告の遅延や市場の価格発見の遅れが生じることを発見した。さらに、Ni et al. (2023) はサステナビリティ・レポートを対象に、GPT による TCFD 適合スコアを開発している。これらの研究は、生成 AI が従来測定することが困難であった事象を抽出または数量化できることを示している。

5 要約と展望

本稿では、AI 技術に関連する会計学の論文の発見事項を体系的に整理した。とりわけ、(1) 機械学習を利用した財務数値の予測に関する分析、(2) テキスト分析を中心とする自然言語処理の応用、そして (3) 生成 AI の出現が会計研究にもたらす新潮流という 3 つの観点から、AI 技術が会計学の分析方法、分析内容および対象範囲をいかに変化させてきたかを明らかにした。

第 2 節で見たように、機械学習の導入は、従来の計量経済学的手法が抱えていた変数制約や推計方法の問題を克服し、予測精度を大きく向上させた。不正会計・倒産・将来業績の予測といった伝統的領域において、機械学習は多数の説明変数を同時に扱うことを可能にし、会計情報の有用性を新たな観点から再検証する基盤を提供した。また、XAI の導入によって、予測モデルのブラック・ボックスの問題が緩和され、説明変数の重要度を確認することを通じて得られる会計的知見が豊富になった。これにより、機械学習を用いる会計学研究は「予測精度」と「理論的解釈」の両立という新しい視点を持つようになっている。

第 3 節のテキスト分析研究は、AI が会計学の対象を「数値」から「言語」へ拡張したことを示していた。企業の開示文書やカンファレンス・コールに含まれる記述情報を、自然言語処理を通じて定量化することにより、経営者の意図やリスク認識を把握し、市場反応や業績予測を説明できることが分った。また、数値情報とテキスト情報を融合した分析や、テキスト分析を用いた企業特性の抽出といった分析も行われており、財務数値の分析が主であった会計研究は大きく展開していることが確認できた。

最後の第 4 節では、生成 AI の登場が会計研究に与えた影響について要約した。初期の研究は AI の会計に関する知識や回答能力を試していたが、近年では、生成 AI を実際の会計関連文書に適用し、その具体的な能力を検証する論文が増えている。生成 AI が有する能力は、テキストの「予測」「分類」「要約」「抽出・定量化」といった要素に分類された。生成 AI が研究手法として定着するかは現時点

では未知数であるが、生成 AI が生み出すアウトカムの有用性を示す経験的証拠が蓄積されれば、生成 AI は研究者にとって新しい分析方法となり、その結果は様々な判断のベンチマークとなる可能性を秘めている。

【注】

¹ AI と会計学研究の関係について議論した論文として、首藤（2019）、首藤（2022）および伊藤・首藤（2024）がある。本稿の第 2 節と 3 節の内容は首藤（2019, 2022）、第 4 節の内容は伊藤・首藤（2024）と関連が深い。本稿の追加的な貢献は、最新の論文を追加したことに加えて、AI に関するすべての研究領域の体系化と展開を要約したことにある。

【引用文献】

- 伊藤広大・首藤昭信, 2024, 「生成 AI と会計学研究」, 『會計』第 206 卷 3 号, 224-237 頁.
- 首藤昭信, 2019, 「AI が会計学研究に与える影響」, 『會計』第 195 卷第 2 号, 15-29 頁.
- 首藤昭信, 2022, 「機械学習が会計学の研究手法へ与えた影響」, 日本会計研究学会特別委員会（代表 奥村雅史）『情報関連技術の進展と財務会計に関する研究 最終報告書』, 251-267 頁.
- Altman, E. I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23(4), 589-609.
- Bao, Y., Ke, B., Li, B., Yu, Y. J., Zhang, J., 2020. Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *Journal of Accounting Research* 58(1), 199-235.
- Barth, M. E., Li, K., McClure, C. G., 2023. Evolution in value relevance of accounting information. *The Accounting Review* 98(1), 1-28.
- Basu, S., Ma, X., Briscoe-Tran, H., 2022. Measuring multidimensional investment opportunity sets with 10-K text. *The Accounting Review* 97(1), 51-73.
- Bernard, D., Blankespoor, E., de Kok, T., Toynbee, S., 2025. Using GPT to measure business complexity (SSRN Scholarly Paper No. 4480309). Social Science Research Network.
- Bertomeu, J., Cheynel, E., Floyd, E., Pan, W., 2021. Using machine learning to detect misstatements. *Review of Accounting Studies* 26(2), 468-519.
- Bertomeu, J., Lin, Y., Liu, Y., Ni, Z., 2025. The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy. *Journal of Accounting and Economics* 80(1), 101782.
- Bilinski, P., 2024. The usefulness of ChatGPT for textual analysis of annual reports (SSRN Scholarly Paper No. 4723503). Social Science Research Network.
- Blankespoor, E., Hendricks, B. E., Miller, G. S., 2023. The pitch: Managers' disclosure choice during initial public offering roadshows. *The Accounting Review* 98(2), 1-29.
- Breitung, C., Müller, S., 2025. Global business networks. *Journal of Financial Economics* 166, 104007.
- Brown, N. C., Crowley, R. M., Elliott, W. B., 2020. What are you saying? Using topic to detect financial misreporting. *Journal of Accounting Research* 58(1), 237-291.
- Bryan, S. H., 1997. Incremental information content of required disclosures contained in management discussion and analysis. *The Accounting Review* 72(2), 285-301.
- Cao, Y., Chen, L., Tucker, J. W., Wan, C., 2025. Can generative AI help identify peer firms? *Review of Accounting Studies* 1-43.

- Cheng, X. (Joyce), Dunn, R., Holt, T., Inger, K., Jenkins, J. G., Jones, J., Long, J. H., Loraas, T., Mathis, M., Stanley, J., Wood, D. A., 2024. Artificial intelligence's capabilities, limitations, and impact on accounting education: Investigating ChatGPT's performance on educational accounting cases. *Issues in Accounting Education* 39(2), 23-47.
- Cheng, Q., Lin, P., Zhao, Y., 2025. Does generative AI facilitate investor trading? Early evidence from ChatGPT outages. *Journal of Accounting and Economics* 101821.
- Choi, G. Y., Kim, A., 2024. Firm-level tax audits: A generative AI-based measurement. *Chicago Booth Research Paper* (23-23).
- Dechow, P. M., Ge, W., Larson, C. R., Sloan, R. G., 2011. Predicting material accounting misstatements. *Contemporary Accounting Research* 28(1), 17-82.
- de Kok, T., 2025. ChatGPT for textual analysis? How to use generative LLMs in accounting research. *Management Science* 71(9), 7888-7906.
- Donovan, J., Jennings, J., Koharki, K., Lee, J., 2021. Measuring credit risk using qualitative disclosure. *Review of Accounting Studies* 26(2), 815-863.
- Drake, M. S., Moon, J. R., Jr., Warren, J. D., 2024. Classifying forecasts. *The Accounting Review* 99(6), 129-156.
- Easton, P. D., Kapons, M. M., Monahan, S. J., Schütt, H. H., Weisbrod, E. H., 2024. Forecasting earnings using k-nearest neighbors. *The Accounting Review* 99(3), 115-140.
- Eulerich, M., Sanatizadeh, A., Vakilzadeh, H., Wood, D. A., 2024. Is it all hype? ChatGPT's performance and disruptive potential in the accounting and auditing industries. *Review of Accounting Studies* 29(3), 2318-2349.
- Guenther, D. A., Peterson, K., Searcy, J., Williams, B. M., 2023. How useful are tax disclosures in predicting effective tax rates? A machine learning approach. *The Accounting Review* 98(5), 297-322.
- Hope, O.-K., Wang, C., Wu, Y., Zhang, M., 2025. Does convergence with international standards on auditing improve audit quality? *The Accounting Review* 100(2), 189-218.
- Jha, M., Qian, J., Weber, M., Yang, B., 2024. ChatGPT and corporate policies (No. 32161). *National Bureau of Economic Research*.
- Jones, S., Moser, W. J., Wieland, M. M., 2023. Machine learning and the prediction of changes in profitability. *Contemporary Accounting Research* 40(4), 2643-2672.
- Kim, A. G., Nikolaev, V. V., 2024. Context-based interpretation of financial information. *Journal of Accounting Research*, forthcoming.
- Kim, A., Muhn, M., Nikolaev, V., 2024. Financial statement analysis with large language models (No. arXiv:2407.17866; Version 2). *arXiv*.
- Kim, A., Muhn, M., Nikolaev, V., 2025a. Bloated disclosures: Can ChatGPT help investors process information? (No. arXiv:2306.10224; Version 4). *arXiv*.
- Kim, A., Muhn, M., Nikolaev, V., 2025b. From transcripts to insights: Uncovering corporate risks using generative AI (No. arXiv:2310.17721; Version 2). *arXiv*.
- Li, F., 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45(2-3), 221-247.

- Li, F., 2010. The information content of forward-looking statements in corporate filings—A naïve bayesian machine learning approach. *Journal of Accounting Research* 48(5), 1049-1102.
- Li, E., Shen, M., Tu, Z., & Zhou, D., 2025. The Promise and Peril of Generative AI: Evidence from GPT as Sell-Side Analysts (No. arXiv:2412.01069; Version 2). arXiv.
- Lin, Y., Shen, R., Wang, J., Julia Yu, Y., 2024. Global evolution of environmental and social disclosure in annual reports. *Journal of Accounting Research* 62(5), 1941-1988.
- Lopez-Lira, A., Tang, Y., 2024. Can ChatGPT forecast stock price movements? Return predictability and large language models (No. arXiv:2304.07619; Version 5). arXiv.
- Lu, Y., Nikolaev, V. V., 2022. Expected loan loss provisioning: An empirical model. *The Accounting Review* 97(7), 319-346.
- Ni, J., Bingler, J., Colesanti-Senni, C., Kraus, M., Gostlow, G., Schimanski, T., Stammbach, D., Vaghefi, S. A., Wang, Q., Webersinke, N., Wekhof, T., Yu, T., Leippold, M., 2023. CHATREPORT: Democratizing sustainability disclosure analysis through LLM-based tools (No. arXiv:2307.15770; Version 2). arXiv.
- Ohlson, J. A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18(1), 109-131.
- Perols, J., 2011. Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *AUDITING: A Journal of Practice & Theory* 30(2), 19-50.
- Purda, L., Skillicorn, D., 2015. Accounting variables, deception, and a bag of words: Assessing the tools of fraud detection. *Contemporary Accounting Research* 32(3), 1193-1223.
- Song, S., 2021. The informational value of segment data disaggregated by underlying industry: Evidence from the textual features of business descriptions. *The Accounting Review* 96(6), 361-396.
- Starica, C., Marton, J. P., 2025. Identifying the relationship between earnings and prices. *The Accounting Review* 100(2), 383-420.
- Wood, D. A., Achhpilia, M. P., Adams, M. T., Aghazadeh, S., Akinyele, K., Akpan, M., Allee, K. D., Allen, A. M., Almer, E. D., Ames, D., Arity, V., Barr-Pulliam, D., Basoglu, K. A., Belnap, A., Bentley, J. W., Berg, T., Berglund, N. R., Berry, E., Bhandari, A., ... Zoet, E., 2023. The ChatGPT artificial intelligence chatbot: How well does it answer accounting assessment questions? *Issues in Accounting Education* 38(4), 81-108.
- Zhang Parker, C. (Abigail), Jiang, L., Cho, S., Vasarhelyi, M. A., 2025. Predicting material misstatements using machine learning. *The Accounting Review* 1-38.