

2025 年 7 月 30 日 全 7 頁

AI とビッグデータ分析、社会科学の新潮流

国際学会に見る ESG・ジェンダー・SNS 研究の最前線

経済調査部 研究員 高須 百華

[要約]

- 2025 年 7 月 2 日～4 日まで、ローマで開催された CARMA (International Conference on Advanced Research Methods and Analysis) と呼ばれる学会に参加した。同学会では、経済学など社会科学に関連するビッグデータや AI を応用した研究発表が行われていた。
- 社会科学分野において AI や機械学習等を用いた研究は学術分野に限らず、実務に応用されることも多い。政策立案や企業戦略における実務的な応用を意識した研究も多く見られ、発展途上の分野に対し、様々な切り口の分析が見受けられた。
- 広範なトピックに関する発表が行われていたが、特に重点的に話されていたのは、ビジネス需要の大きい ESG の分野と、ホットトピックであるジェンダー、また社会的に問題が大きくなってきている SNS の研究であった。今回の学会では、これらのテーマに対して、自然言語処理 (NLP) をはじめとする、AI や機械学習の先端技術を用いた分析手法が積極的に導入されており、社会科学分野における新たな研究の潮流が形成されつつある。

はじめに

2025 年 7 月 2 日～4 日まで、筆者はローマで開催された CARMA (International Conference on Advanced Research Methods and Analysis) と呼ばれる学会 (以下、学会) に参加した¹。経済学など社会科学に関連するビッグデータや AI を応用した研究発表が行われ、トピックは多岐にわたった。内容は、AI やテキストマイニング、ESG、金融、政策、労働経済、需要予測、計量経済学、Web・SNS 分析など幅広く、現代社会経済の課題に対する多角的なアプローチが見られた。

会場がローマだったこともあり、イタリア語の自然言語処理 (NLP) の研究や、欧州全域に焦点を当てた研究が多かった。イタリアをはじめ欧州では、日本でも顕在化しつつある社会経済の長期的な不確実性や停滞、SNS 上での過激な言説や分極化、それに伴う政治・社会の分断といった課題が、研究の根底にある問題意識として共有されていたように思う。欧州で記録的熱波が発生している中の学会参加であったが、ローマでも森林火災が発生しており、ESG に関する発

¹ 7th International Conference on Advanced Research Methods and Analytics website, [CARMA 2025](#)

表では参加者の質問や議論は活発だった。欧州で特に、ESG 推進をはじめとする企業や個人の責任ある行動が強く求められている背景には、このような甚大な自然災害による被害や、排外主義などによる社会の分断、戦争が身近な問題として発生していることが関係しているのだろう。

社会科学（経済学など）の分野では、SNS の普及やそれに伴う社会の複雑化、データの多様化から従来使用されてきた伝統的な分析手法では実態が把握できず、研究が追いつかなくなっている。こうした状況を受けて、新たな研究手法の導入や、分野を跨いだ連携の在り方が模索されているところだ。特に、AI や機械学習、NLP といった技術を活用したアプローチが注目されている。また、学会では政策立案や企業戦略における実務的な応用を意識した研究も多く見られた。実際に企業や中央銀行、省庁などで働いている人の研究発表も多く、発展途上の分野に対し、様々な切り口の分析が見受けられた。今回は学会で発表されていたトピックのうち、特に発表が多かった ESG、ジェンダー、SNS の 3 つについてまとめている。

自然言語処理や機械学習による定量的な ESG 評価、その推進や批判

本学会では学者だけではなく、企業のエンジニアなども多く参加しており、企業と大学の共同研究の発表なども行われていた。企業の ESG 関連の取り組みに関する発表では、企業がどのように ESG 情報を収集・評価しているのかに耳目が集まっていた。投資家にとって非財務情報の重要性が高まる中、企業による情報開示も増加傾向にあるが、その内容は多様な文書形式や情報媒体で散らばっており、これらを比較可能な標準形式にデータとして整理、つまり構造化することが近年の課題となっている。今学会で発表された研究の多くは、手法については概ね同様のアプローチを採用していた。Web スクレイピングや PDF からのテキスト情報の抽出といったように、NLP を用いてテキストデータを順序尺度²や数値データに変換、比較可能な情報として活用するというものであった。

EU では企業サステナビリティ報告司令（CSRD、Corporate Sustainability Reporting Directive）と、欧州サステナビリティ報告基準（ESRS、European Sustainability Reporting Standards）によって、サステナビリティ情報開示を義務付ける制度と基準が定められている。ただし、中小企業を中心にこれらの基準を満たすことが難しい企業も多く、情報の開示媒体も統一されていないので、何ををもってどのように評価するかということに関しては発展途上の段階にある。

ESG 関連の企業活動は、単純に環境（Environment）、社会（Social）、ガバナンス（Governance）という枠組みにとどまらず、より広範なトピックをカバーしており、企業ごとに独自の取り組みを行っている可能性があるということが、学会の研究で示唆されていた³。こうした多様性もあり、ESG に関する金融情報のテキストデータを収集し、大規模言語モデル（LLM）に学習させ

² データを順位として示し、「1: 良い、2: ふつう、3: 悪い」など段階で評価する場合などをいう。

³ 先行研究として下記の文献が挙げられていた。

Bronzini M., C. Nicolini, B. Lepri, A. Passerini & J. Staiano (2024). “[Glitter or gold? Deriving structured insights from sustainability reports via large language models](#),” *EPJ Data Science*, 13, 41

るための金融情報特化型コーパス（大量の構造化された文章データベース）の構築を試みる研究もあった⁴。データの収集対象は、年次報告書や開示情報に限らず、ニュースや規制当局への提出書類など、できる限り多様な情報源に広がっており、企業の ESG 活動をより正確に把握するための基盤づくりが進められている。

ただし、機械学習が出した評価のみで判断すると信頼性に欠ける可能性もあるため、難しい手法を使用せずに、従来の方法で人が行った評価と機械学習が作成した評価指標の比較分析を行った研究もあった。その結果、人が作った評価と比べると機械学習の作成した指標ではバイアスが排除できていないことから、機械学習による新しい ESG 評価方法は可能でも、現状は人が行った評価との併用が望ましいという結果が報告されていた。

一方、ESG 推進に対する批判的な視点を示す研究も発表されており、EU 圏では環境問題への関心が薄れてきているのではないかという指摘もされていた。また、SDGs 推進による Green Transition のコストの大部分を企業が負担している現状が、雇用に悪影響を及ぼしているのではないかという研究もあった。これについては機械学習を用いて、地方の失業率の予測が行われており、従来の手法では用いられてこなかった、環境保護の強度や大気汚染物質の排出量などのデータが用いられていた。この結果、機械学習を用いたいくつかのモデルは高い予測精度を誇り、失業率予測は線形にはならずある閾値を超えると、環境保護の強さと大気汚染物質の排出量の多さの両方が失業率を引き上げる傾向が見られたと報告されていた。

ジェンダー関連の分析でも広がる AI 活用

社会科学分野のビッグデータ等の応用では他に、ジェンダー関連の研究も多く発表されていた。上述の ESG の内容とも関連し、企業関連や投資、またジェンダー指標と SNS の文脈等での研究発表が行われていた。ESG 関連の研究と同様に、近年はジェンダー関連の研究でも NLP を用いて、企業のウェブサイトや報告書、政策文書、SNS 投稿など多様な情報媒体に掲載されたテキストデータを分析する手法が広がっている。これらは、従来の定量的な指標等では捉えきれない企業の非財務的な側面や社会的文脈を読み解く新たなアプローチとして注目されている。また、NLP 以外の機械学習等の手法を用いた分析も盛んに行われており、分野横断的な数値データを基に新しい指標を作る研究等もあった。

従来型の研究としては、イタリアのライフサイエンス企業がジェンダー平等に関する発信をしているかどうかを分析した研究があった。企業規模や収益よりも、女性 CEO の存在の有無がジェンダー平等の情報発信の有無と相関しているという結果が出ていた。女性 CEO を選出する企業はジェンダー平等にも高い関心を持ち、企業文化や価値観としてそれを積極的に発信する傾向があると考えられる。

また、EU では EIGE (European Institute for Gender Equality) という組織がジェンダー平

⁴ なお、今回の学会で発表された論文は内容を見ることができないものが非常に多いため、本稿では参考文献の紹介を省略している。実際に発表された論文のタイトルや著者を確認したい場合は、脚注 1 のウェブサイトに掲載されている CARMA のプログラムを参照されたい。

等指標等を作成し、国レベルのジェンダー平等の状況をモニタリングしている。これを基に最新の手法を用いて、地域レベルのジェンダー指標の作成を試みた研究もあった。具体的には、変数間の因果関係を含む構造を学習できるベイジアンネットワーク⁵という確率モデルを用いてジェンダー平等指数を作成していた。国レベルの指標と同様の手法により地域レベルでジェンダー平等指数を作成しようとしても、データが不足していることも多い。この欠損データを従来の単純な補完手法で対応するのは難しかったが、ベイジアンネットワークを用いた推計により、不完全なデータを効果的に扱うことが可能となっている。また、シナリオ分析も行っており、例えば「どの変数に介入すればジェンダー平等指数を効果的に改善できるか」や「教育参加率の変化によってどの程度ジェンダー平等指数が改善するか」など、政策提言における本手法の有用性を示していた。

Keynote Speaker であった英シェフィールド大学の Diana Maynard 氏は、SNS をはじめ、Web 空間におけるジェンダー差別に基づく誹謗中傷が、表現の自由にとって世界的に重要な問題になっていると述べていた。多くは公言しないものの、女性著名人は何らかの言論の暴力にさらされており、特にジャーナリストの 20%程度が言論による暴力を受けているとされていた。これらの Web 空間上のジェンダー差別は言論によるものだけではなく、ミームの作成や画像の拡散、また、Web 上に限らず、本人を含め家族や近い人が実際に何らかの被害を受ける場合も多い。そこで、Maynard 教授は X、YouTube のコメント欄、テレグラムや Instagram などの SNS、ニュースサイトへのコメント欄などのデータを取集し、特定の人物に対する誹謗中傷をリアルタイムで収集・分析、分類して可視化するダッシュボードを International Center for Journalists (ICFJ) と共同で開発している⁶。SNS の誹謗中傷の拡散の震源地となっている人物や、それに近い人をウォッチリストに加えることもでき、危険に対するアラートとなっている。ただし、テキストベースのデータの集積、分析、分類等を行っているので、画像を用いた言語以外の誹謗中傷、多言語やスラングなどへの対応は困難だとされていた。Maynard 教授は、今後は全世界的に、SNS でのふるまいに対する責任のある行動の促進と、公権力による規制や法整備の強化が求められると述べていた。

影響力が高まる SNS 上の発言や反応をどのように分析するのか

近年、SNS の分析の重要性が増している。元来、公共での言論空間は新聞やテレビなどの情報媒体によって形成されてきたが、SNS の普及により、個人が容易に情報を発信・共有できるようになったことで、言論の場は大きく変容している。特に X や Instagram、YouTube、TikTok といったプラットフォームでは、特定の話題が瞬時に拡散され、世論形成や政策決定にまで影響を及ぼすケースも少なくない。このような背景から、SNS 上の発言や反応を定量的・定性的に分析

⁵ 本研究では、オブジェクト指向ベイジアンネットワーク (OBN) を適用していた。この手法は、部分ネットワークをクラスとして定義・カプセル化し、そのインスタンス (オブジェクト) を組み合わせてシステム全体を構築する。このアプローチにより、複雑な問題を独立した部品の集合として階層的にモデル化できるため、モデルの再利用性、拡張性、そして保守性を大幅に向上させることが可能となる。

⁶ ICFJ, [Online Violence Alert and Response System](#)

することは、社会の動向を把握する上で極めて有効である。例えば、災害時の情報伝達、排外主義的な発言の構造把握、政治的な支持動向の可視化など、応用範囲は多岐にわたる。

災害時の情報収集や緊急対応において、SNS 上の発言を迅速かつ的確に分析する技術の重要性が増している。特に、災害発生時には膨大な量の投稿が瞬時に拡散されるため、緊急対応機関が必要な情報を効率的に抽出することが求められる。そこで、分析にかかる時間やコストを大幅に抑えた小型言語モデルを活用し、災害関連情報の識別精度を高めることを目的とする研究もあった。限界として、火災など特定の語が他の文脈でも頻繁に使われることから誤分類が起りやすいなど実践的な課題も挙げられていた。

また、SNS 上の発言を単なるテキストとして捉えるのではなく、発言の構造や相互作用、発言に対するポジティブ・ネガティブなどの感情の流れまでを含めたネットワーク的な視点からの分析が注目されている。この文脈からイタリアの移民問題に対しての 2018-2022 年の約 2,000 万件の Twitter（現在の X）上のデータを用いて、移民関連ワードの投稿と Google Trends との相関を調査し、社会的な関心の定量評価を試みた研究があった。実際、Twitter の投稿量と Google Trends の検索量は高い相関を示し、移民関連の政治的イベントのタイミングで明確に投稿・検索が増加していた。また、分析の結果、「人道的支援」と「反移民措置」に分極したクラスタが可視化されていた。

2024 年のイタリアで行われた欧州議会選挙に向け、主要政党リーダーの Instagram 投稿へのコメントを対象にコメント間の相互作用構造と、ポジティブ・ネガティブなどの感情の時系列変化を分析した研究もあった⁷。2024 年時点の右翼側リーダーに対する言論空間のネットワークは広く、ある程度一元化していた。一方、左翼側ネットワークでは決まった人物の相互交流が多く、広がりやが少なく、加えて内部で対立しているということが分かった。また、上位アクティブユーザー50 名によるネットワークは双方向の会話が多く、高い密度を示し活発な政治討論を反映していた。選挙が近づくにつれ感情の極性が高まって、対立的な言説が増加し、ポジティブな共感よりもネガティブな対立の影響力が強いことが明らかにされていた。選挙期においては、限られた数のハブユーザーが議論の主導権を握り、情報の拡散・世論操作に関与している可能性があり、ネガティブな対立を緩和する介入策の検討などが必要となると述べられていた。

さらに、Keynote Speaker の Maynard 氏によって、情報のサプライチェーン（Information supply chain）という概念（**図表 1**）が紹介されていた。適切な情報のサプライチェーンは民主主義に必要なものであるが、現在この全ての段階において誤情報や攻撃の危機にさらされている（関連する文献として、例えば Seger et al. [2025]⁸がある）。原因として、コミュニケーション

⁷ 主な先行研究の例として下記の文献が挙げられていた。

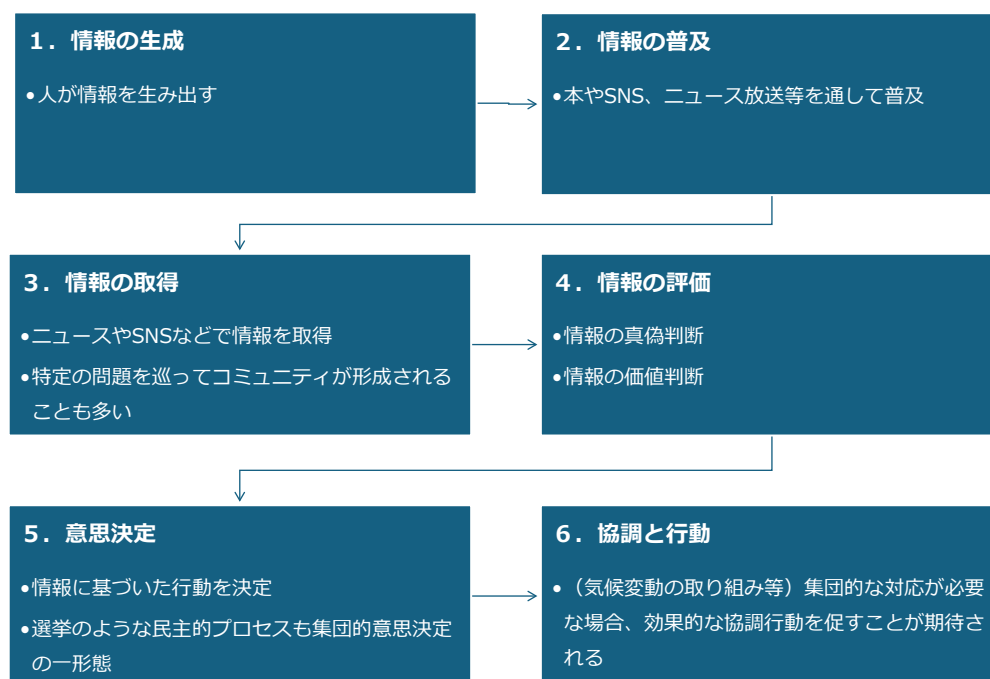
Arman, Z. R. (2024). “[Network Issue Agenda Setting on Facebook: Exploring the Interplay Between Polarized Campaigns and Party Supporters](#),” *Social Science Computer Review*.

Parmelee, J. H., S. C. Perkins, & B. Beasley, (2023). “[Personalization of politicians on Instagram: what Generation Z wants to see in political posts](#),” *Information, Communication & Society*, 26(9), 1773-1788

⁸ Seger, E., H. Perry, & J. Hancock (2025) “[Epistemic Security 2029 : Fortifying the UK’ s information supply chain to tackle the democratic emergency](#),” DEMOS.

オンラインツールの急速なデジタル化（情報の分散と集中が同時に進行）、報道のエコシステムの弱体化（従来型のメディアの影響力の低下）、外国からの影響や規制の不備などが挙げられていた。正しい情報が適切に情報の受け手に届かず、不信感や陰謀論などが市民の間で拡大し、世界各地で起こっているポピュリスト政党の台頭や、民主主義の後退、政治的な分断の拡大に繋がっている。SNS を通じて情報が瞬く間に拡散する一方、それに対する透明性や説明責任の確保が欠けた環境にある。情報のサプライチェーンの各段階の脆弱性を点検し、情報のサプライチェーンの健全化・強化が国家的な取り組みとして必要になっていると述べていた。

図表 1：情報のサプライチェーン



（出所）Seger et al. (2025) より大和総研作成

まとめ

今回、社会科学分野（経済学など）での AI 等新技術の応用に関して、特に重点的に話されていたのは、ビジネス需要の大きい ESG の分野と、ホットトピックであるジェンダー、また社会的に問題が大きくなってきている SNS の研究であった。今回の学会では、これらのテーマに対して、NLP をはじめとする、AI や機械学習などの先端技術を用いた分析手法が積極的に導入されており、社会科学分野における新たな研究の潮流が形成されつつあることを強く感じた。

ESG に関しては、企業の情報開示の多様性とデータの構造化、評価手法の課題など、実務と理論の両面から多角的な議論が展開されていた。ジェンダー関連では、新たなジェンダー指数の構築や SNS 上の差別構造などの研究が発表されていた。SNS 分析においては、災害対応や政治的分極、感情の流れの構造化など、社会の動態をリアルタイムで把握するための技術的アプローチが試みられていた。

これらの研究は、単なる技術応用にとどまらず、社会の複雑性に対する理解を深め、政策や企業行動、個人の意識などを把握し、影響を与える可能性を秘めている。今後、社会科学分野においては、こうした技術と理論の融合による実証的かつ応用的な研究がますます重要になるだろう。今学会は、まさにそのような未来への方向性を示す場であり、今後の調査活動においても大きな示唆を与えるものであった。