

近未来金融システム創造プログラム第8回講義レポート

第8回目となる本日は、筑波大学システム情報系社会工学域教授、独立行政法人経済産業研究所上席研究員（特任）、マネーフォワード総合研究所シニアフェローの小西葉子氏を迎え、「答えはデータの中にある」という題目で講義が行われた。公的統計と民間データを組み合わせた社会課題の捉え方、能登半島地震後の復興、災害時消費、生成AIの利用実態、Well-beingの測定についてお話しいただいた。

社会の解像度を高めるデータの見方

どの分野でも、データを使った可視化や統計解析は必要になる。欧米やシンガポールなどではデータビジュアライゼーションが独立した講義として設けられている一方、日本では統計学や計量経済学を学んでも、最後にグラフをどう描くかまで体系的に学ぶ機会はまだまだ限られている。グラフを描く際には、軸の取り方に加え、時系列や地域といった方向、粒度や周期をどう設定するかによって、データの特徴の見え方が変わる。

2012年頃のビッグデータブーム以降、POS、人流、衛星データなどを政策立案や公的統計に活用する動きが広がった。足元では生成AIによって研究と教育の環境も変わりつつあり、文献検索、壁打ち、グラフの試作、PythonやRのコード生成などを通じ、研究と教育の作法そのものを変えつつある。

データを見るとき、出発点になるのは「なぜこのグラフはこの形になるのか」という違和感である。データを入手したらまずグラフにし、上に伸びるスパイク、下に落ちるスパイク、安定している部分を確認する。さらに、時系列や地域といった方向からばらつきを見たり、前年同週比で平時からどれだけ離れているのかを見たりすることで、単なる増減の背後にある構造を考えていく。

例えば防災では、家具固定が推奨されていても賃貸住宅の居住者が取り組みにくい理由や、防災リュックの中身が季節や地域を問わず同じでよいのが問題になる。米不足の局面では、米を備蓄しているのか食べているのかを考えるため、ふりかけの売上に着目する発想が出てくる。こうした違和感を起点に、生活者視点の問いを具体化していく。

能登復興と公的統計の役割

能登半島地震後の復興を考えると、社会の解像度を高めるデータの見方が具体的な課題として浮かび上がる。能登地方の純付加価値額は2,723億円で石川県内の11.3%、加賀地方は2兆1,274億円で88.7%を占める。半島という地理的条件と経済規模の小ささが重なる地域で災害が起きたとき、復旧・復興が遅れやすい構造をどう測るかが課題になる。

令和3年経済センサス活動調査を用いると、全事業所・全業種を対象に、能登と加賀を分けた産業構成を見ることができる。能登地方では製造業の比率が約29%で、全国の

16.4%、石川県全体の 24.7%、加賀地方の 24.1%より高い。復興を「元に戻す」こととして捉えるなら、製造業の回復はベンチマークになりうる。さらに県外本社比率を見ると、奥能登では地場産業の比重が高く、資本や人の接点が相対的に少ない構造も見えてくる。

民間ビッグデータは速報性や詳細性に優れる一方、企業が自社事業のために取得するため、地域や対象に偏りが生じやすい。能登の先端部には民間 POS データの対象店舗が十分に入っていない場合があり、被災者アンケートも実施しにくい。このような状況では、生産動態統計調査や商業動態統計調査のような政府統計は、民間データでは捉えにくい地域も含めて把握する手段になる。

民間データと政府統計は、どちらか一方に置き換えられる関係ではない。民間データは足元の変化を細かく見るために役立つが、その結果がどの程度社会全体を表しているのかを考えるには、網羅性の高い公的統計による答え合わせが必要になる。政府統計を弱めてしまうと、民間データの結果を検証しにくくなり、国際比較の基盤も失われかねない。生成 AI や民間ビッグデータの利用が広がる場面でも、公的統計は民間データの結果を検証する基盤として必要になる。

災害時消費と生活リズムの可視化

災害時の消費行動は、災害の種類と季節によって異なる。能登半島地震のように突発的で局所的な災害と、南海トラフ地震臨時情報（巨大地震注意）のように広域へ警戒を促す災害では、人々の備え方や購買の広がりが変わる。水の販売動向を前年同週比で見ると、能登半島地震では北陸で急激な上昇が見られ、南海トラフ地震臨時情報では東海地方だけでなく全国にも反応が広がる。

災害への備えは、水のように多くの場面で必要になるものだけでなく、地域や季節によって必要性が変わるものも含めて考える必要がある。2024 年 5 月に実施された約 21,000 人へのアンケートでは、携帯カイロを備えている割合が、日本列島を北から南へ見るにつれて低くなる傾向が見られた。命の安全が確保された後の段階では、冬であればカイロ、夏であれば制汗剤や冷感用品のように、避難生活の快適性を支える品目が必要になりやすい。防災を考える際には、災害の直後に必要なものと、その後の生活を支えるものを分けて捉える視点も求められる。

生活者の行動変化は、POS データからも読み取ることができる。POS データでは、「売れているもの」だけでなく「売れていないもの」に注目することで、人々の行動の変化が見えてくる。コロナ禍では、マスクや手指消毒剤の売上増に加え、日焼け止め、口紅、ファンデーションなど、マスクで隠れる部分に使う化粧品の売上減が、外出自粛やマスク着用の広がりをうかがわせた。販売額をそのまま比べると一枚の図に収めにくい場合でも、品目ごとの販売額順位に置き換えることで、2019 年と 2020 年の変化を比較しやすくなる。さらに、2019 年から 2022 年までの週次順位を円環状に並べるランククロックを用いると、季節ごとの売れ方や、コロナ禍を含む期間中の変化を一枚の図で捉えやすくなる。

生活や企業活動の変化を記録するという点では、生成 AI の利用実態も重要な対象になる。生成 AI については、企業ログと個人調査の二つのデータから利用状況を把握できる。企業側では、マネーフォワードのクラウド会計データを用い、約 87,000 社の支払いログから導入時期や業種差を見る。個人側では、2026 年 1 月にインテージリサーチと実施した全国アンケートにより、約 51,000 人を対象に、利用頻度、利用場面、AI に任せる度合いを把握している。

企業ログでは、2022 年 11 月の ChatGPT（GPT-3.5 ベース）の公開を契機に、2023 年初頭から独立型生成 AI サービスへの支払いが立ち上がる様子が見える。従業員規模で切ると大きな違いは見えにくい一方、業種で切ると ICT や教育サービスが相対的に高く、飲食、不動産、建設、運輸などは低い側に位置する。生成 AI は業務全体へ一律に浸透するというより、既存のデジタル基盤や、それぞれの業務で使いやすい場面に応じて、選択的に利用されているように見える。

個人調査では、生成 AI の使い方を、自分で考える際の参考情報として使う「参照」、人と AI が作業を分担する「協働」、AI に作業を大きく任せる「委任」に分けている。文章作成、要約・構成、プログラミング、金融資産形成、健康相談、悩み事など、調査設計時に想定しきれない使い方でも自由記述に現れていた。インターネットや PC の普及初期には、どこからどのように使われ始めたのかを示すデータが十分に残らなかった。生成 AI についても、無償版の高度化や既存ソフトへの組み込みが進むと、独立したサービスとしての導入経路や利用状況が見えにくくなる。その前に、企業と個人の両面から利用過程を記録する意義がある。

Well-being をどう測る？ 主観と客観の接続

Well-being は、GDP だけでは捉えにくい豊かさを考える指標として、政策の場でも身近になりつつある。内閣府の「満足度・生活の質に関する調査」では、生活の満足度を 0 から 10 までの 11 段階で尋ね、健康や社会とのつながりなど複数の領域についても調査している。一方で、Well-being、主観的満足度、幸福度といった言葉は近い意味で使われることも多く、何を測っているのかを丁寧に見る必要がある。

満足度データを見ると、女性が男性より高い、高齢者が高い、都市部が高いといった傾向が見られる。これらは Well-being のデータではよく現れる傾向だが、賃金や就労機会、健康不安、物価といった別の政策課題と並べると、一見矛盾して見える。こうした違和感は、満足度の源泉を考える入口になる。就労や賃金の課題と結びつけて語られやすい女性の満足度がなぜ高いのか、健康不安が高まりやすい高齢者の満足度がなぜ高いのか、物価の高い都市部の満足度がなぜ高いのかを考えるには、主観的な満足度だけでなく、暮らしぶり、時間の使い方、消費、地域のアメニティなどの客観データと接続して見ることが求められる。

満足度の源泉を考える視点は、地域の Well-being にも広げられる。地域については、そ

こに住んでいる人と、観光や寄付などを通じて外から関わる人の視点を分けて考えることができる。居住者以外の人を惹きつけるものを持つ地域では、居住者の Well-being も高いのではないかという仮説を立てることができる。観光、ふるさと納税、備え力は、地域の魅力や外部との関係を測る代理変数になりうる。

地域の魅力を測る際には、病院、温泉、ホテルの数のように、施設の有無や件数だけでは年ごとの変化を追いきにくい指標も多い。そのため、インバウンド旅行者数や国内旅行者数のように、実際に人が地域へ動いているかを示す人流データが材料になる。外国人旅行者の動きは、キャッシュレス化や Wi-Fi 環境、交通インフラなどの進み具合を考える材料にもなる。一方、地域間移動が止まると観光データだけでは分析が難しくなるため、移動せずに地域と関係を持つふるさと納税のデータも有効な材料になる。自治体ごとの寄付額や件数は長期に追うことができ、観光、ふるさと納税、備え力といった市町村レベルの客観データを組み合わせることで、Well-being の地域差の説明や新たな地域指数の作成につながる可能性がある。

金融に近い例として、2026 年 1 月の制度利用調査から、ふるさと納税、確定拠出年金、新 NISA、iDeCo、旧 NISA などの利用状況と Well-being の関係を見ることが出来る。ふるさと納税の利用者は Well-being が高く、女性の方が男性より高い傾向も見られる。ただし、ふるさと納税の利用が満足度を高めていると単純に読むことはできない。所得との相関や逆の因果も考えられるほか、制度を利用する人の属性や、新しい制度・時代の変化への関心が、制度利用と満足度の双方に関係している可能性もある。

主観的な満足度は、ある瞬間を切り取る写真のような性格を持つ。調査で聞かれた時点の気分や場面に左右される一方、毎秒のように細かく測ればよいというものでもない。満足度を生活の実態に近づけるには、どの程度の頻度で、どの場面を捉えるのかを考える必要がある。24 時間という共通の時間の中で、朝起きてから寝るまでの行動、時間の使い方、消費の内容、その人にとって大事な場면을満足度と結びつけられれば、総合的な満足度がどのような生活の積み重ねから形づくられているのかを考えやすくなる。家計調査や家計消費状況調査に満足度の項目が加われば、日々の消費や高額な商品・サービスの購入と満足度の関係も見えやすくなる。

生成 AI 時代にデータと統計学・データサイエンスをどう使うか

生成 AI 時代には、グラフの試作やコード作成が容易になり、研究や教育の進め方も変わりつつある。それでも、違和感や社会課題をどのように問いにするのか、どの粒度や周期でデータを見るのか、データ自身の構造から社会の構造をどう考えるのかは、引き続き問われる。写真や映像が時代の記録であるように、グラフも人々の生活の積み重ねと社会の変化を記録する手段になりうる。

Q&A

Q1. インフレ下で外国人料金やデュアルプライスを導入すべきかを、満足度データによって示すことはできるのか。

A1.

小西（敬称略）：デュアルプライスについては、報道では個別店舗へのインタビューが中心で、国の会議などで十分に議論されている印象は強くない。導入の是非や対象、既に導入している国の状況などを机上調査し、500 人から 1,000 人規模のアンケートも実施できれば、影響を考える材料になりうる。

Q2. 仮説と合わない結果が出た場合、集めたデータをどのように見極め、取捨選択しているのか。

A2.

小西：基本的には、自分よりもデータを優先する姿勢をとっている。社会から一時的にデータを借りているという感覚があり、データがそこまでしか言わないならそこで止める。都合の悪い観測値を理由をつけて切り落とし、結論を強く見せることはしない。一方で、データの質や経験値が上がるにつれ、事前の仮説と大きく外れることは少なくなっている。

Q3. 仮説を立てる力やセンスは、どのように高められるのか。

A3.

小西：仮説を持ってデータを確認する姿勢が大事である。計算結果は誰が計算しても同じだが、そこからどのようなインプリケーションを出すかは、センスというより経験に左右される。読んで、書いて、自分の領域を考え続けることに加え、自分の感覚を疑い、異なるレイヤーの人や現場に触れることが、社会的な仮説やビジネス上の仮説を立てる助けになる。

Q4. グラフから変数同士の関係を見る際、見せかけの相関や誤った解釈を避けるにはどのような視点が必要か。

A4.

小西：グラフを見る際には、線、散布図、地図など、限られた型をその都度組み合わせている。相関を見る際には、逆の因果、同時発生、X 軸側を外生的に扱えるかに注意する。一方で、母集団に近いビッグデータが手に入る場合には、相関そのものも出発点として有用になりうる。

Q5. 能登地方の県外本社比率の低さから、どのように仮説を立て、検証し、論理的な結論につなげたのか。

A5.

小西：能登の分析は、データを見てからではなく、「なぜ復興が遅いのか」という問いから始まった。東日本大震災では企業や人の動きが見えた一方、能登ではそれが見えにくかったため、地域を気にかける人や企業との関係を測る方法を考えた。経済センサスで見られる事業所のネットワークを数えることで、縁を持つ人や企業が相対的に少ないことが、震災時には不利に働きうるという見方につながった。

Q6. 個人の生成 AI 利用ログを分析することは制度的に可能なのか。制度上難しい場合、それを使えるようにする活動はあるのか。

A6.

小西：生成 AI のログについては、Google トレンドのように粗い粒度で利用の動きが見えるようになる可能性はある。ただし、個人を細かく追うことには慎重であり、人は自由に変わってよい存在である。個人の履歴を細かく分析できるようにする方向というより、社会の変化を捉えられる程度に集計された、粗い粒度のデータを見る方向が考えられる。

Q7. 政策決定において、Well-being のような指標はシンプルに使うべきか、それとも複数指標を組み合わせて構造まで見るべきか。

A7.

小西：Well-being を各省庁のアウトカムや KPI にすることは慎重に考えるべきである。満足度が著しく低く、政策やセーフティーネットが届いていない可能性を知る用途には意味がある。一方で、Well-being は場面によって変わり、子どもと大人でも捉え方が異なりうるため、一定の丁寧さも必要になる。ただし、複雑にしすぎると自治体や省庁が再現できず、ブラックボックス化しやすい。政策指標として使うには、実生活に即した丁寧さと、継続して計算できるシンプルさの両方が求められる。

赤井（敬称略）：Well-being にはオブジェクティブなものとサブジェクティブなものがある。前者に属する客観的なデータの束で後者の主観的な満足度を擦り合わせながら表現していく、ということが政策の評価の分野では始まっている。