

ビッグデータの“料理”の仕方

近年、ビッグデータという言葉をよく耳にします。私たちの知らない間に大量のデータが収集されています。一方、心理学では目的に応じて意図的にデータを収集し分析します。ここでは、心理学内外の研究におけるビッグデータの“料理”の仕方の相違点や、その誤りによって生じる問題を掘り下げます。（坂田陽子）

ビッグデータと心理学 ——自然主義的心理学は可能か？

専修大学人間科学部心理学科 教授

国里愛彦

部屋の窓から穏やかな陽の光が差し込み、真っ青な空が見える。気持ちの良い午後だ。

私たちは何らかの窓を通して世界から情報を得ている。現代では、スマートフォン（以下スマホ）のディスプレイが多いだろうか。私たちがスマホのディスプレイを覗いて世界の情報をあれやこれや得ているのと同時に、スマホを通して私たちの足跡が残る。私がどこにいるのか、見たものや得た情報について何を呟いたのか、誰にメッセージを送ったのか、スマートウォッチを使っている場合は活動量、心拍や体温、睡眠のリズムなどの情報も足跡として残る。スマホを超えて、コンビニなどの買い物、通院などの情報も私たちの足跡として蓄積されている。これらの情報は私たちの心そのものではないかもしれないが、心がたどった跡になる。

ビッグデータと心理学

ビッグデータに明確な定義はないものの、心理学でビッグデータに言

及する場合は、何らかの研究目的のもとで収集されたものでなく、自動的に蓄積された大規模なデータという意味で扱われる。SNSのログなど、大量の情報を解析する場合などは、ビッグデータを用いた研究と言えるだろう。このような研究は計算社会科学¹といった研究領域としてまとめられ、有益な知見を提供している。臨床的な研究で例を挙げると、国によっては認知行動療法を実施した際にその前後のアセスメント結果を症例登録する制度があり、他のデータベースと連結されていたりする。一般的に心理療法の効果を検証する上では無作為化比較試験が行われるが、このような症例登録の情報を活用し、因果推論の手法などを用いることで、無作為化比較試験をしなくとも妥当な有効性判断をすることもできる²。

ビッグデータの魅力は、研究者が特定の目的を持ってデータ収集を行っているのではなく、自動的にデータが蓄積されており、その数の力で事象間の関連や実態を明らかにできることであ

る。その一方で、事前に目的に合わせて計画され、厳密に統制されて測定されたデータではないので、研究者の目的と合致するとは限らないこと、十分な統制ができておらず結果の妥当性に疑問が残る場合もある。また、プライバシーなどへの配慮も必要なため、個人を深く知ろうとする研究には向かないなどの限界もある。何十万人のデータから得られる平均的な心ではどこか物足りない。かといってビッグデータを踏まえると、1人から細かに深掘りした心では、それが他の人にも当てはまるのか確信を持ってない。ビッグデータを前に、こういうジレンマを感じるかもしれない。これはビッグデータに限らない話であり、古くから心理学において議論されてきていることだろう。そこで、少し視点を変えて、それが測定される状況の自然さに焦点を当てたい。

自然な状況で残る足跡

今、穏やかな陽の光を感じるリビングのクッションで、私は寝転びながら

この文章をスマホに打ち込んでいる。実にリラックスした姿勢ではあるが、何を書いたものか悩み、時にSNSでしようもない情報を見たり、またスマホに文章を書き込んだりしている(図1)。これは心理学の研究において、心理学実験室に参加者を呼んで実験をする状況とは随分と違って、実に自然な状況と言えるだろう。ビッグデータとして集められる情報にはいろいろとあるだろうが、スマホなどを通して得られる情報は、生きている人が自然な状況で活動している様が足跡として残ったものになる。ビッグデータには、データが大規模であるというメリットだけでなく、できるだけ生の状態の人間の心を扱うことができるかもしれないという側面がある。

図1 今日も足跡を残し続ける



自然主義的心理学は可能か?

心理学においては実験や調査をする場合に、文脈の情報などをできるだけ統制し、抽象化された材料を用いることが多い。そのようなデータ収集をすることで、小さなサンプルサイズであっても何らかの結論が得られるように工夫されてきた歴史がある。そ

の一方で、ビッグデータを用いた研究は、そのような統制され抽象化された方法では、自然な状況で私たちの心が芽生え、動く様子を見逃してしまっているのではないかという疑問を突きつけているように思われる。そのような疑問は、私の専門とする計算論的精神医学においても投げかけられ、*naturalistic computational psychiatry*を考慮すべきという議論もある³。*naturalistic*をどう訳したものか悩ましいが、ここでは自然主義としておこう。心の病を検討する上では、できるだけ文脈の影響を統制した認知課題を準備し、必要に応じて脳画像の測定をするなどがこれまでの研究方法だった。それをもっと自然に近づけていく努力をしようというのが論旨になる。

心理学において、抽象化しシンプルにした実験課題や状況の設定は、心が存在する環境をモデル化した結果だろう。それによって状況に依存しない本質を知ることができるとも考えられるが、同時に一般化可能性の問題も招き入れることにもなった。ビッグデータを用いた心理学研究はそのデータの大きさよりも、自然主義的な方向への窓を開くことに意義があるのかもしれない。現在、深層学習や大規模言語モデルを活用した心理学研究も進んできている。これらの機械学習モデルは私たちが生活する自然な環境を定量化し、それを研究に組み込むことを可能にする。これまで抽象化しシンプルにせざるを得なかったが、これらの技術によって、そのままの自然な刺激や状況における人間の反応を研究することが可能になりつつある。このような流れを受けて、自然主義計算論的認知科学という研究領域も提案されている⁴。ビッグデータそして大規

模言語モデルなどの機械学習モデルの活用は、心理学の新たな領域を開拓し始めている。果たして、自然主義心理学は可能だろうか?

何十万人のデータから得られる平均的な心では心をつかんだというにはどこか物足りない。一方、1人を徹底的に深掘りして得られる心は何かを私たちに伝えてくれるがそれが他の人に当てはまるのか分からない。ビッグデータをめぐる議論は、私たちが取り組む心理学研究が平均的な因果関係を同定することだけでいいのか、個人の中での心の動きを記述することが必要なのではないかを改めて考えさせる。また、理想化しシンプルに整理された実験状況や調査によって得られたデータからは分かりやすい結論が得られるが、それが私たちの自然な心なのだろうか。一方、私たちが生活するのと同じくらい複雑で自然な状況で得られたデータは自然な心を扱っていると思われるが、それを適切に扱う方法は定まっているとは言い難い。ビッグデータをめぐる議論は、私たちの心を測定したり扱う上で、自然主義を取り入れることができるのかという問題を突きつける。願わくば、リビングのクッションの上でゴロゴロしながら窓の外やスマホを眺めながらも移ろいゆく私の心を扱えるような、そんな方法論の発展がビッグデータや大規模言語モデル・深層学習の発展とともに形作られることを期待している。皆様はどうでしょう?



くにさと・よしひこ

2011年、広島大学大学院医歯薬学総合研究科創生医科学専攻修了。博士(医学)。2021年より現職。専門は臨床心理学、計算論的精神医学、認知行動療法。著書に『R/Pythonではじめる計算論的精神医学』(共著、金芳堂)など。

1 Lazer, D. et al. (2009) *Science*, 323(5915), 721-723. 2 Rosenström, T. H. et al. (2025) *Lancet Psychiatry*, 12(3), 189-197. 3 Palaniyappan, L. et al. (2025) *J Psychiatry Neurosci*, 50, E67-E72. 4 Carvalho, W., & Lampinen, A. (2025) *arXiv*, 2502.20349. *COI:本記事に関連して開示すべき利益相反はない。