

計算社会科学研究のための大規模データベース・ データ解析・シミュレーション拠点の形成

① ビジョンの概要

ビッグデータ、機械学習に代表されるデータ解析技術、高度なシミュレーション技術等は近年急速に発展しているが、必ずしも技術的知識を有さない人文社会科学系研究者にとっては、容易に活用できるものではない。これらの技術が社会経済現象に関わる学術研究に円滑に活用され、我が国における計算社会科学研究、さらには異分野融合研究が飛躍的に活性化する研究体制の構築が本構想のビジョンである。

② ビジョンの内容

我が国の自然科学分野では、巨大な観測機や実験設備等を共同利用し共同研究を行うことにより、ノーベル賞に繋がるような革新的な成果をあげてきたが、人文社会科学分野においては同様の取り組みは乏しい。近年、人文社会科学系の文部科学省「共同利用・共同研究拠点」を中心に資料やデータの共同利用・共同研究が活発に進められているものの、比較的小規模に留まっており、自然科学分野のような国際的に顕著な成果には至っていない。しかし、人文社会科学分野においても、自然科学分野のように大規模な観測機や実験設備等を共同利用し共同研究を行うことで、ノーベル賞クラスの国際的に顕著な研究成果をあげることができるとは思えないだろうか。

現代社会においては、社会の大規模観測機が実質的に実現しつつある。パソコン、タブレット、スマートフォン等のデバイスを用いたオンライン上での人々の行動やコミュニケーションは日々デジタルで記録され、大規模なソーシャルデータが急激な勢いで蓄積されている。これらのビッグデータの蓄積とともに、機械学習に代表されるデータ解析技術も近年急速に発展している。さらには、スーパーコンピュータに代表される計算技術の発展に伴い、現実社会を再現できるレベルの大規模シミュレーション技術も発展してきた。

このような大規模データと計算技術を用いて社会経済現象を分析するのが「計算社会科学」と呼ばれる新しい研究領域である。計算社会科学は、データサイエンス、計算科学、および社会科学に跨る文理融合領域であり、近年急速に発展している。しかし、社会科学であるにも関わらず、同領域の研究を行っているのは情報工学系の研究者が多く、人文社会科学系の研究者は比較的少数派である。その主な原因は、必ずしも技術的知識を有さない人文社会科学系の研究者にとっては、大規模データ、機械学習、シミュレーション等の先端技術は容易にアクセスできないからであると考えられる。

人文社会科学分野における大規模データや大規模シミュレーション技術の活用は、自然科学分野における大規模観測機や実験装置の共同利用に匹敵するものである。これらのデータや技術を人文社会科学分野において十分に活用することができれば、我が国における計算社会科学研究のみならず、文理融合を中心とした多種多様な異分野融合研究が飛躍的に活性化するのはないか。

以上の問題意識のもと、本ビジョンでは20年後の学術戦略の目標として以下を掲げる。

1. 人文社会科学系を含む様々な分野の研究者が必要とする社会・経済に関するデータを、研究者の要望に応じて、必要があれば新たに収集し、可能な限り提供するための設備体制と人的体制の構築
2. 必ずしも技術的知識を有さない人文社会科学系を含む様々な分野の研究者が、多種多様なデータと、現時点では機械学習に代表されるような高度なデータ解析技術を円滑に活用するための設備体制と人的体制の構築
3. 必ずしも技術的知識を有さない人文社会科学系を含む様々な分野の研究者が、現実的な社会実験を仮想的に行うためのシミュレーション・システムの構築と、同システムを円滑に活用するための設備体制と人的体制の構築

③ 学術研究構想の名称

計算社会科学研究のための大規模データベース・データ解析・シミュレーション拠点の形成

④ 学術研究構想の概要

本構想と既存の共同利用機関等との主な違いは、既存の共同利用機関等では既に整備されたデータ等が提供されているのに対して、本構想では、研究者からの要望により、必要に応じて新たにデータを収集し、研究者が活用しやすい形でデータを整備・提供する点である。データ解析技術とシミュレーション・システム

の共同利用に関しても、研究者からの要望に応じて、研究者が活用しやすい形で提供することが本構想の特徴である。

⑤ 学術的な意義

計算社会科学は社会科学であるにも関わらず、同領域の研究を行っているのは情報工学系の研究者が多く、人文社会科学系の研究者は比較的少数派である。その主な原因は、必ずしも技術的知識を有さない人文社会科学系の研究者にとっては、大規模データ、機械学習、シミュレーション等の先端技術は容易にアクセスできないからであると考えられる。本構想の目的は、人文社会科学系を含む様々な研究者が大規模データ、機械学習、シミュレーション等の先端技術に容易にアクセスできないという障害を克服し、計算社会科学研究、さらには文理融合を中心とした多種多様な異分野融合研究を飛躍的に活性化させることである。SNS や大規模データの分析等は比較的歴史が浅く、我が国において計算社会科学研究を強力に推進することにより、学術的重要性の高い同領域において世界的イニシアティブを握ることができる可能性がある。国際的には、計算社会科学においては人種、差別、犯罪等に関する研究も盛んに行われているため、我が国において解決不可能と思われる様々な社会問題に対しても、新たな視点から解決策が提示される可能性がある。計算社会科学は社会科学、データサイエンス、計算科学の融合領域であるため、多くの分野と密接に関わっており、特に社会学、経済学、政治学、統計学、情報工学、人工知能、計算科学とは繋がりが深く、これらを含む様々な関連分野への波及効果が期待できる。

⑥ 国内外の研究動向と当該構想の位置付け

「計算社会科学」(Computational Social Science) という呼称は、2009年に学術雑誌「Science」に掲載された同タイトルの論文により広く知られるようになった。計算社会科学は欧米を中心に急速に発展を続けているが、我が国においても「計算社会科学研究会」が2016年に発足し、2021年には「計算社会科学学会」として正式に学会となった。同年には、同学会のメンバーが中心となり、我が国初の計算社会科学のテキスト「計算社会科学入門」(丸善出版)を出版した。本構想の中心機関である神戸大学計算社会科学センターは、2018年に我が国初の計算社会科学に特化した研究センターとして設立され、同年に世界初の計算社会科学に特化した国際学術雑誌「Journal of Computational Social Science」をSpringer社より創刊した。同誌は、2023年にインパクトファクター3.2が付与され、国際的に高く評価されており、同センターおよび本構想は、国内外の計算社会科学研究において中心的な役割を果たしていると位置づけられる。



図 設備・人的体制

⑦ 社会的価値

計算社会科学に対する国民の関心は高く、フェイクニュースの拡散など、日常的にも関心の高い話題が研究対象とされている。また、大規模データから見えてくる社会構造に対する知的関心は高く、その知的価値も高い。さらに、大規模データ・機械学習・シミュレーション等による経済予測等は、産業的価値も高く、効率的なSDGsへの取り組み方を提示する等により、SDGsへも貢献できる。

⑧ 実施計画等について

実施開始前年度に連携機関との連携体制に関する協議等を行い、実施1年目には、若手研究者、設備管理技術者、事務補佐員、URAの採用、データ収集・整備専用サーバの導入、深層学習用GPUワークステーションの導入を行う。実施2年目以降は以下を順次行う：シミュレーション用スーパーコンピュータの導入、ハイブリッドストレージシステムの導入、大規模データの収集・整備に関する要望の公募と共同利用、データ解析技術に関する要望の公募と共同利用、汎用的社会経済シミュレーション・システムの完成、汎用的社会経済シミュレーション・システムに関する要望の公募と共同利用。

⑨ 連絡先

上東 貴志 (神戸大学計算社会科学センター)