

社会科学の多角的統計情報データアーカイブ構築とエビデンス・ベースド・ポリシー・メイキング (EBPM) の実現：21 世紀の社会科学の創造に向けて

① 計画の概要

本計画は、日本における本格的な社会科学のデータアーカイブの構築とそれを用いた各種の社会経済政策や民間事業の推進に寄与する仕組みを提案するものである。21 世紀を特徴づけるものは情報通信技術を軸とした産業革命であり、統計情報をインターネットで集め、各種の行動分析やマーケティングに生かしてビジネスを行ってきた Google, Amazon, Facebook and Apple (GAFA) を中心とした企業がそれを牽引してきた。翻って、日本は 1990 年代より、バブル経済の後始末に終始し、産業革命を牽引するような企業や発想を生み出すことができないで来た。これが失われた 30 年の根本的な原因である。この間、アメリカを中心に、ヨーロッパ諸国もインターネットを通して集められる各種の情報処理のための技術革新に全力を尽くし、多くの有益な技術的ブレークスルーを生み出して来た。

本計画では、日本政府の行って来た各種の政府統計調査のデータアーカイブのデジタル化および複数の政府統計をマッチングさせて使うためのデータ・インフラ整備などに加えて、国勢調査や経済センサスなどの全国レベルの悉皆調査による母集団情報（地域メッシュ統計）等の整備および提供を目指すものである。

政府を中心とする政治経済の領域でも欧米先進国では、政府統計データの二次利用の促進、政府統計と民間統計データを組み合わせた商業利用、業務統計である税務データの公開による税制分析など急激な進展が見られた。他方、日本では政府統計を産業革命のエンジンとして使うという発想もなければ、民間事業に政府統計を公開するという英断も出来なかった。また、学術利用もその分野や目的が極めて限定的であった。これらの反省に基づいて、本計画では本格的な社会科学のデータアーカイブを構築し、広く社会の共同利用共同研究拠点として、それを提供して行くことを目的とする。

② 学術的な意義

社会科学データアーカイブが構築され、それが世界に真に開かれた共同利用共同研究拠点として機能することになれば、これまで日本の社会科学統計が一部の研究者にしか利用できなかった制度の壁、言語の壁が一挙に取り除かれることになり、社会科学の実証研究の裾野が大幅に拡大することが期待できる。実際、日本経済は世界の「課題先進国」である。少子高齢化、デフレ経済、長期経済停滞、過剰な政府債務、原発問題、地震津波などの自然災害など挙げたらきりが無い。これらの問題が統計的にどのように記録されており、日本社会はこれらの問題にどのように対処してきたのか、あるいは対処しそこなってきたのかを統計的に確認できれば極めて貴重な情報となる。しかもこれらの問題は世界にとっても重要な課題として認識されており、その問題解決策は普遍的な価値を持っている。

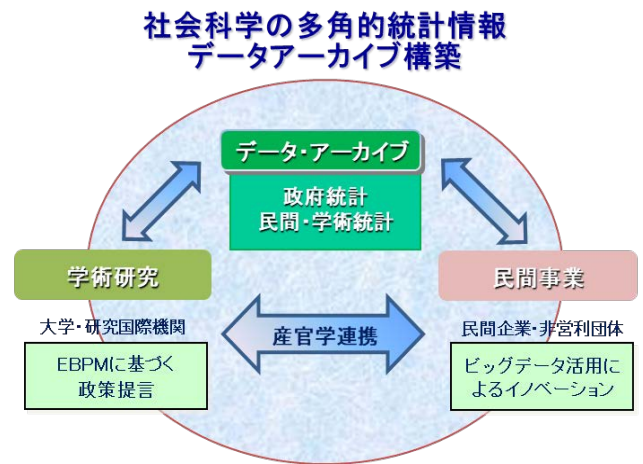
社会科学は、様々な社会現象や時々刻々発生してくる社会経済問題を多角的な統計情報データによって分析することを重視している。そのために必要になるデータを、様々な社会経済現象を多様な観点から定点観察しておくことが不可欠であり、それら観察により解析が可能になるものである。すなわち、統計データの収集は経済学、社会学、政治学など多岐にわたり、しかも、その調査項目もかなり広範なものが必要とされる。多くの場合、同一分野でも補完的な複数の調査が必要になる。ここに融合領域として社会科学のデータアーカイブ構築の必要性があり、複数の研究分野の統計情報データの収集を行う価値があると考えられる。

③ 国内外の動向と当該研究計画の位置づけ

国際的に見て、社会科学のデータアーカイブとして、社会調査データやマクロ経済の基礎データの収集・保管・活用を行ってきたのは、1962 年から活動を始めたミシガン大学の社会調査研究所 (Institute for Social Research) およびそれに附属する Inter-university Consortium for Political and Social Research (ICPSR) である。また、イギリスのエセックス大学内にある UK Data Archive も 1967 年に社会科学研究機構データバンクとして発足し、2000 年代に入って、政府統計の委託データアーカイブとして機能している。この他ヨーロッパのみならずアジアでも多くの社会科学データアーカイブが設立されている中で、日本には政府統計やそれに準じるマクロ経済の基礎データを収集・公開する機関が欠けてきた。今後のデータ競争の時代に充実した社会科学データアーカイブを持たずに、様々なレベルでの国際競争ができるとは到底考えられないし、国際的な留学生の獲得競争にも影響がでてくるだろう。

④ 実施機関と実施体制

本計画は、日本における本格的な社会科学のデータアーカイブの構築とそれを用いた各種の社会経済政策や民間事業の推進に寄与する仕組みを提案するものである。実施機関の長は一橋大学学長である。一橋大学は社会科学の総合大学として、我が国の中で重要な役割を担うだけの実績と人材、組織力をもっている。具体的に本計画を担うのは、一橋大学経済研究所附属社会科



学統計情報研究センターである。前述の通り、1964年4月の日本学術会議の勧告に基づき、人文社会科学の全国共同利用施設「一橋大学経済研究所附属日本経済統計文献センター」として発足、2002年4月に「社会科学統計情報研究センター」への拡充改組し、社会科学統計情報の収集に取り組むとともに、新たにマイクロデータ分析セクションを設け、マイクロ統計の利用を促進する活動を行ってきた。当センターは、2006年度には財団法人日本統計協会統計活動奨励賞を受賞し、2007年9月には日本統計学会第3回統計活動賞を受賞するなど、広く学界コミュニティに認知されている。

本計画では、マイクロデータ提供を先行して行ってきた東京大学社会科学研究所附属社会調査・データアーカイブ研究センター、慶應義塾大学経済研究所パネルデータ設計・解析センター、大阪商業大学JGSS研究センターと連携し、社会科学データアーカイブの構築とその情報提供のノウハウを共有し、インフラ構築の礎とする。また、社会経済政策のための実証研究やその基礎研究に関しては、長年共同研究を行ってきた連携機関である、京都大学経済研究所、大阪大学社会経済研究所、神戸大学経済経営研究所との連携を考えている。これらの機関との連携については双方の合意が得られている。

⑤ 所要経費

総経費 70 億円 (10 年計画)

人件費

シニアレベル研究者 1000 万円×10 名=1 億円、ポストドクレベル研究者 700 万円×10 名=7000 万円、技術職員 300 万円×10 名=3000 万円、合計 2 億円/年

調査費

マイクロデータ収集 (パネル調査・社会調査を含む) 7000 万円×4 件=2.8 億円、データベース形成 (デジタル化・政府統計加工等) 1 億円、デジタル・アーカイブ・システム構築 1 億円 (第1-5年度)、合計 4.8 億円/年 (1-5年) 3.8 億円/年 (6-10年)

設備費

高性能ネットワークサーバー システム一式 1 億円 (5年で更新: 初年度+第6年度)、通常ネットワークコンピュータ、プリンター、計算ソフト等 5000 万円、合計 1.5 億円/年 (1年、6年) 5000 万円/年 2-5年、7-10年)

初年度 8 億 3000 万円、2-5年度 7 億 3000 万円、6年度 7 億 3000 万円、7-10年度 6 億 3000 万円、合計 70 億円 (10 年計画)

⑥ 年次計画

初年度—5年度:

(1) 社会科学データアーカイブの基本構造を構築する。具体的には、技術的問題の洗い出し、暗号化技術、異質なデータ情報の格納および連結方法などの検討。

(2) 社会科学データアーカイブに収納するデータの選択とそれに向けての政府との交渉。

(3) 民間保有ビッグデータと社会科学データアーカイブにある政府母集団情報とのマッチング。実用可能性に向けての試行。

(4) ミクロ調査 (パネルデータ、社会調査) を用いた政府統計調査への情報補完。調査相互のバイアスチェック (高度統計人材の育成)。

(5) 国際データアーカイブと国内データのマッチングに向けたデータ整備。

6年度—10年度:

(6) データアーカイブ・システムが構築されることを前提に、国際的データアーカイブへのデータ格納を開始。

(7) データアーカイブを用いて、EBPM に向けた研究および講習の開始。(政府向け講習会の開催)。

(8) データアーカイブの改善およびシステム改良。

(9) データアーカイブを用いたビジネスへの利用分野の開拓 (利用収益を見通した実用化)。

(10) データアーカイブでの雇用システムの確立と人材の確保。

計画期間終了後には、海外で先行しているミシガン大学社会調査研究所やエセックス大学 UK Data Archive などの例に倣って、利用料や講習料の徴収などを通して、社会科学データアーカイブが持続可能となる仕組みを構築する。また、ビジネスへの利用機会の拡大を通して、ビジネス界からの支援・協賛金なども得ることを考えている。それらの資金と一橋大学および連携機関の協力を得て、日本および世界の統計を担う人材の育成と供給を図っていく。

⑦ 社会的価値

21 世紀はデータの世紀と言われ、社会経済の多くの活動や機能がインターネットを通して行われるようになってきた。いまやネット上を行き交う統計情報は経済活動の原動力であると言ってもいいだろう。政府統計の不適切処理が問題になり、統計調査の精度の向上や品質の管理の必要性が広く認知されるようになったが、政府統計は莫大な予算をかけて、定期的に収集している貴重な公共財であり、その二次利用の促進は、時代の要請と言えよう。さらに、政府統計のデジタルデータおよびその調査票情報、集計表などが公開され、一般利用に資することになる経済的・産業的価値はかなり大きい。民間統計にとっては、多くの政府統計は、貴重な母集団情報を提供してくれるものであり、政府統計を民間統計で代替することはほぼ不可能である。政府統計を中心とした社会科学データアーカイブの構築は、単に学界コミュニティに限らず、国際機関、中央政府、地方自治体、民間企業など幅広い対象に大きな貢献をするものと期待できる。

⑧ 本計画に関する連絡先

北村 行伸 (一橋大学・経済研究所)