

Title	行政データの活用と課題：EBPMの推進に向けて
Author(s)	小林, 淑恵; 榊原, 裕二
Citation	年次学術大会講演要旨集, 33: 280-283
Issue Date	2018-10-27
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/15667
Rights	本著作物は研究・イノベーション学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Research Policy and Innovation Management.
Description	一般講演要旨

2 A 0 8 行政データの活用と課題～EBPMの推進に向けて～

小林淑恵（文部科学省），榊原裕二（NISTEP）

yoshietk@nifty.com

1. 背景

行政機関では、政策立案、また行政施策の推進を目的とし、統計法によらない情報収集として様々な調査が実施されており、いわゆる業務統計と呼ばれている。これまで全国的な大規模な調査であっても当初の業務のための情報を得た後に再利用されることはまず無かった。しかしこれらの情報は学術研究において非常に重要であり、エビデンスベースの政策立案において積極的に分析、活用することが期待されている。現在の、オープンデータの流れもこのデータの活用促進を後押ししている。本報告では、教育・科学技術分野の行政データの活用、公開について最新の情報を紹介するとともに、その課題等を明らかにする。

2. 内容

日本の公的統計は、2007年の統計法改正以降、大きく「基幹統計」と「一般統計」に分けられるようになった。基幹統計は国勢調査、国民経済計算、家計調査、労働力調査等、国の行政機関が作成する統計のうち総務大臣が指定する特に重要な統計であり、現在56種ある。一般統計は、公的統計のうち基幹統計以外の統計である。これ以外にも「業務調査・業務統計」と呼ばれるものが、統計法に寄らない調査、統計として数多く実施され、行政における業務の中で活用されている。業務統計の結果は公表されるが、e-statにも掲載されず、広く行政担当者や研究者に活用されることは基本的でない。

しかしオープンデータの潮流の中、いくつかのデータの再利用、2次利用が試みられ、本報告では、その例として科学技術・学術政策研究所『博士人材追跡調査』と文部科学省 初等中等教育局『全国学力・学習状況調査』の2つを紹介する。

2.1. 科学技術・学術政策研究所『博士人材追跡調査』

博士人材追跡調査とは

科学技術・学術政策研究所では、日本の博士人材の就業状況を雇用統計に即した形で継続的に把握するために、『博士人材追跡調査（Japan Doctoral Human Resource Profiling, JD-Pro）』を平成26年から開始した調査である。初回のコホートは平成24年度に博士課程を修了者、2回目のコホートは平成27年博士課程修了者であり、それぞれ有効回答者数は5千名前後、回収率は約4割弱であった。

この1次集計の結果は、以下の報告書にまとめられ、科学技術・学術政策研究所から公表されている。
＜1次集計＞

・『博士人材追跡調査』第1次報告書-2012年度博士課程修了者コホート 文部科学省科学技術・学術政策研究所 NISTEP REPORT No.165(2015.11)

・『博士人材追跡調査』第2次報告書 2017.8.24

このローデータを詳しく検証した2次分析には、現在までに公表されたものに以下のような物がある。今後、さらにデータの活用が進めば、多くの論文等に活用されることになろう。

＜2次分析＞

小林淑恵「女性博士のキャリアと家族形成」文部科学省科学技術・学術政策研究所 DISCUSSION PAPER No.147 (2017.6)

小林淑恵「博士の入職経路とマッチング効果」文部科学省科学技術・学術政策研究所 DISCUSSION PAPER No.148 (2017.6)

柴山創太郎・小林淑恵「博士課程での研究指導状況とインパクトー「博士人材追跡調査」による総合的な分析ー」文部科学省科学技術・学術政策研究所 DISCUSSION PAPER No.150(2017.7)

Shibayama, Sotaro and Yoshie Kobayashi“Impact of PhD Training: A Comprehensive Analysis based on a Japanese National Doctoral Survey”Scientometrics (2017.8)

本報告では、博士のキャリアパスに関する最新の分析（榊原 2018）を紹介する。

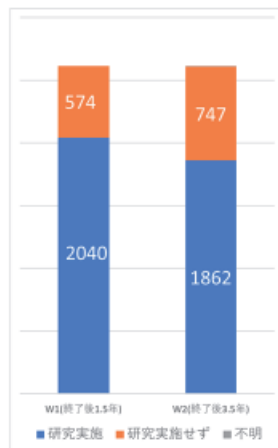
・「博士人材追跡調査」を使った分析の可能性

博士人材追跡調査は過去 2 回の調査を個票ベースで連結して分析できる。ここに着目し、ストックベースの変化量だけでなく、フローベースの変化量を得ることが出来るので、実際のキャリアパスに近い情報を得ることが出来る。下記に一例を示すが、ストックベースでは、僅かな変化しか認められないが、フローベースでは大きなクロスする流れがあることが一目瞭然である。

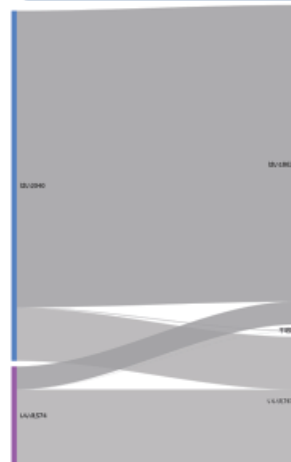
「博士人材追跡調査」の分析事例（検討中の分析の手法）

「博士人材追跡調査」はパネル調査なのでストック情報だけでなくフロー情報を抽出が可能。

フロー毎の母集団を分析することにより高精度の分析が可能。



	W1(終了後1.5年)	W2(終了後3.5年)
研究実施	2040	1862
研究実施せず	574	747
不明		5



W2(終了後3.5年)			
	研究実施	研究実施せず	不明
W1(終了後1.5年)	1729	307	4
	133	440	1
	1862	747	5

「具体的な研究活動を行っているか」(W1Q46)と「研究を行っていますか。」(W2Q22)の集計結果

単純集計

個票で紐付け

さらに、フロー毎の母集団に着目し、母集団の特徴を抽出することにより、キャリアパスを形成する因子についての示唆が得られるものと考え、現在分析を進めている。

・「博士人材追跡調査」を使った分析における課題

「博士人材追跡調査」は、まだ 2 回の調査が実施されたばかりであり、今後データの蓄積が必要であることは言うまでもないが、知見を収斂していく上で以下の課題があると考えている。

- ① 「博士人材追跡調査」の結果を実際のキャリア形成過程にいる者に対するフィードバックを考慮する必要がある。解りやすいものでなければならない。
- ② キャリアパスを視覚化し、要因分析をするための手法・知見いずれも不足しており、多くの試行錯誤・検証が必要である。関心を持つ研究者の層を厚くしなければならない。
- ③ 今のところ、キャリアパスに着目した調査は「博士人材追跡調査」以外に存在しない。データの有効利用を考えるべきである。

2.2. 「全国学力・学習状況調査」「全国学力・学習状況調査」

(1) 「全国学力・学習状況調査」とは

全国的な児童生徒の学力や学習状況を把握・分析し、教育施策の成果と課題を検証し、その改善を図るために平成 19 年度から実施しているものである。義務教育の機会均等と、義務教育水準の維持向上

の観点で、継続的な検証改善サイクルを確立し、学校における児童生徒への教育指導の充実や学習状況の改善等に役立てることが目指されている。

(2) データ活用の状況

「全国学力・学習状況調査」の各教科の結果と質問紙調査の結果は、毎年、国立教育政策研究所から公表を行っている。毎年、教科の問題が異なることから、得点は経年比較できない。しかし質問紙に関しては同一の間である場合、経年変化を示すことができ、例えば、「算数（数学）の授業の内容はよくわかりますか」等の問は平成 19 年度から毎年調査しているため、10 年間の間に授業の理解度は増している等の変化を示すことができる。

質問紙調査の自己肯定感や学校の取組等に関しては、教科の正答率とのクロス集計を示し、簡易的ではあるが、相関（効果）について議論するためのエビデンスとして活用されている。

「全国学力・学習状況調査」実施の概要

		本体調査					補完調査			特記事項
		国語・算数 (数学)	理科	英語	児童質問紙	学校質問紙	保護者に 対する調査	経年変化 分析調査	教育委員会に 対する調査	
H19	悉皆	○			○	○				
H20	悉皆	○			○	○				
H21	悉皆	○			○	○				
H22	抽出	○			○	○				
H24	抽出	○	○		○	○				
H25	悉皆	○			○	○	○	○	○	きめ細かい調査
H26	悉皆	○			○	○				
H27	悉皆	○	○		○	○				
H28	悉皆	○			○	○		○		
H29	悉皆	○			○	○	○			
H30	悉皆	○	○		○	○				英語予備調査の実施
H31	悉皆	○		○	○	○				A B 問題統合

(3) データ貸与の開始

データの貸与については、有識者会議で繰り返し議論されてきた。貸与するデータをその活用の用途、申出者の範囲が異なる 3 つのパターン（パブリックユースデータ、匿名データ、個票データ）を準備し、特に個票データに係る貸与のための基本原則、手続き、審査、結果の取扱い等についての指針を示す「「全国学力・学習状況調査」の個票データ等の貸与に係るガイドライン」を平成 29 年 3 月に策定した。

パブリックユースデータについては現在、すでに HP に掲載されており、自由にダウンロードでできるようになっている。また、匿名データの申請受け付けも開始され、こちらは 2 か月に 1 回程度の審査で貸与の承認が出る。個票データ貸与手続きは年に 2 回の申請期間を設けており、第 1 回の申出期間は、平成 30 年 10 月 1 日から末日までである。

「全国学力・学習状況調査」貸与データの種類と概要

データの種類	主な用途	申出者の範囲	データの年度	データの特徴	データサイズ	利用期間	申請	審査
個票データ	1) 学術研究 2) 公的施策の推進	1) 大学、公的機関の研究者 2) 公的機関の役員・職員	調査実施 全年度	申出者の必要最低限のデータ	申出内容による	原則1年	期間中の申請 10月→11月末審査 1月→3月末審査	有識者会議による 個別審査
匿名データ	1) 学術研究 2) 公的施策の推進 3) 大学院生の教育	1) 大学、公的機関の研究者 2) 公的機関の役員・職員 3) 大学共同利用機関	調査実施 全年度	【児童生徒データ】 10%のサンプリング 1) 匿名データの公開可とした市区町村教育委員会に属する児童生徒の情報 2) 公立学校に属する児童生徒 3) 実施日に調査を受けた 【学校データ】 学校規模、地域規模で10%層化抽出	児童・生徒 各年約97,000人 *抽出調査年は約25,000人	原則1年	随時申請	学力調査室による 簡易審査 (2か月に1度)
パブリックユース データ (疑似データ)	目的不問	不問	H27年度 (悉皆、理科)	乱数による疑似データ	学校データ 各100校 児童・生徒 各2,000人	不問	不要	HPからダウンロード

3. 今後の課題

行政の保有するデータの活用が、今後一層の促進することに間違いはない。しかし、データ活用の制度設計、ルール作り、個人情報保護等の観点からデータの効果的な匿名化、より幅広いユーザーのニーズに答える仕組みの構築等、様々な課題がある。その解決策は、「多分野融合」が一つのキーワードであると考えられる。

データ活用の制度設計に関しては2次分析等でデータを貸し出すための個人情報の取扱い、使用できるデータの範囲の決定、手続き等のルール作りには法的な専門家が必要であろう。さらに分析・集計システムの構築に関しては工学系や情報科学系の専門家、あるいはデータサイエンティストの技術が必要だろう。行政機関やそれに準じる組織・機関において、様々な専門性を持つ人々の連携によって、業務データの活用、オープンデータの流れは今後、一層進展させる必要があるだろう。

本報告ではその先駆けとも言える2つの調査の活用について紹介を行った。これに続き、多くの行政データが多くのユーザーに安全に活用され、政策のエビデンスの充実と、これによる効果的な政策の立案、また効果の検証等に活用され、PDCAサイクルの円滑な実施に今後、益々、生かされることを期待したい。