

データアクセスの行方

一橋大学経済研究所

神林龍

2013年3月28日

統計委員会と統計利用者との意見交換会

1. オンサイトとリモートアクセス

2. オンサイトと行政データ

1. オンサイトとリモートアクセス

Workshop on Data Access (WDA)

- 各国政府の統計部局・研究機関の集まり
 - 2007年 Nuremburg (IAB)
 - 2009年 Cardiff (ONS)
 - 2010年 Ann Arbor (ICPSR)
 - 2012年 Luxembourg (EuroStat)
- 現在の先進諸国の政府統計のアクセスの現状の報告・将来の方向への議論

1. オンサイトとリモートアクセス

データアクセス方法は2つの方法に収斂

- Onsite
 - Data research center
 - Remote access
- Public use file
 - Scientific use file
 - Campus use file

1. オンサイトとリモートアクセス

オンサイト施設の各国の違い

- **物理的環境**については各国似たようなもの
 - 管理された空間、ネットワーク的にも孤立した環境
- 主要な違いは**統計部局のチェック**のありよう
 - 原則として計算結果は、統計部局のチェックの後に持ち出し(出力)が認められる。
 - その場でチェックするか(BLS, CB)、「リモートで」チェックするか(CAN, European countries.....)。

1. オンサイトとリモートアクセス

–「リモートで」チェックできる場合

- オンサイトの運営そのものにはそれほど大きな負担はかからない。(オンサイトに常駐する要員の人件費が削減可能)
- オンサイト施設の利点は、物理的に管理された空間のみになる。

⇒オンサイト施設の運営費用をいかに安くあげ、数を増やすかが問われるのは道理

1. オンサイトとリモートアクセス

リモートアクセスの現状

－ オンサイト施設のリモートアクセス化

- ・ オンサイト施設そのものにデータを移すのではなく、中央のサーバーにリモートアクセスすることにより、オンサイトに必要な設備を軽減する。(CAN, FRA, GER)

－ 個別の研究室をオンサイトとみなし、リモートアクセスを認める。(NET, DEN, SWE)

- ・ 原理上、オンサイト施設を国内に限定する必要がない。

2. オンサイトと行政データ

－ オンサイト化の動機

- きっかけは、事業所データの匿名化が困難なこと。
 - 現在では、いくつかの国では、行政データの活用に重点が移っている。
 - － GER、USAにおける雇用保険データ
 - － DEN、SWE、FINなどにおける医療保険データ
- c.f.) Eurostat による調査の一元化

2. オンサイトと行政データ

2種類の行政データ

－ 統計作成に付随するデータ

- 統計ではないという意味で行政データ(統計付随データ)
- 要図、事業所番号対照表、原票など

－ 行政に付随するデータ

- 通常はこちらの意味(行政データ)
- 雇用保険、医療保険、通関、課税記録など

2. オンサイトと行政データ

統計付随データ

－ 統計委員会で扱える範囲

- 早急な保存が必要。
- 統計の事後的検証。
- 産業(職業)分類の再分類など遡って必要になる場合もある。

2. オンサイトと行政データ

行政データ

ー 統計委員会で扱える範囲ではない？

- 設計された調査と異なり、データの構造が複雑で、他データとのリンクが必要など、準備に時間と労力がかかる。

e.x.) 米国における雇用保険データ

- データ構築に関して研究者集団との協力、投資が不可欠

ー 日本の場合、過去の遺産を利用する有力な方法

- そのためには、過去の行政データをアーカイブ化する方法を早期に検討する必要がある。

まとめ

- データを研究者に直接渡すのであれば、Public Use File の形にする
 - 33条のような形での利用は欧米と比較すると例外的
- オンサイト化とリモートアクセスの検討の必要
 - テクノロジーフリーにするためには、リモートアクセスが便利。
- 行政データの保存と整理
 - 統計付随データに関しては統計委員会の範疇のはず