

わが国におけるデータカタログの整備に向けて —現状とこれからの課題—

日本学術振興会 人文学・社会科学データインフラストラクチャー構築推進センター

前田 幸男¹、伊藤 伸介²

日本学術振興会は、2018年11月から「人文学・社会科学データインフラストラクチャー構築事業」を開始した。本事業の課題の1つは、人文学・社会科学の分野横断的なデータカタログの整備にある。こうしたカタログの整備を進めるにあたり、海外のデータアーカイブ機関で展開されてきたデータシェアリングの現状とメタデータ(データに付随する属性情報を表すデータ)の作成状況を把握することは有益と考える。

社会学や政治学の分野では、主として、社会調査データに関する二次分析(secondary analysis)を可能にするために調査データの共有が行われてきた。アメリカでのデータシェアリングについては、1962年にミシガン大学に設立された ICPSR(=Inter-university Consortium for Political and Social Research)に遡ることができる。また、イギリスでは、1967年にエセックス大学に UKDA(=UK Data Archive)が設置され、データシェアリングが展開されてきた。

イギリス最大の人文学・社会科学に関するデジタルデータ(digital data)のアーカイブ施設である UKDA は、社会調査の個票データや政府統計のサーベイデータ等の様々なデータの収集・保管を行っており、UKDS(=UK Data Service)を通じて、7,000以上のデータセットの提供が行われている。イギリスでは、1970年代初期に実証的な社会科学研究が進展したものの、データシェアリングに関する文化的環境の未成熟さ、データの寄託に関する基準の厳しさ等もあり、サーベイデータの収集は困難な状況にあった。さらに、1980年代になると、社会科学研究に対する政府予算が削減されたが、そのことが、イギリスにおける調査個票データの二次分析の促進とデータシェアリングのさらなる受容をもたらした。その後、データアーカイブ機関としての機能を拡充させることによって、UKDA は現在国際的なアーカイブのネットワーク CESSDA(=Consortium of European Social Science Data Archives)における拠点の1つとなっている(UK Data Archive(2007))。

ヨーロッパ諸国には、UKDA の他にも、ドイツの GESIS(=Leibniz-Institute for the Social Sciences)、オランダの DANS(=Data Archiving and Networked Services)、フィンランドの FSD(=Finnish Social Science Data Archive)、ノルウェーの NSD(=Norwegian Centre for Research Data)等、様々なデータアーカイブ施設が存在し、一定の方針に基づいてデータの収集・保管・提供を行っている。また、これらのアーカイブ施設では、それぞれのルールにしたがってメタデータの整備を行っているが、調査・研究データの場合、データ記述に関する国際的な基準である DDI(=Data Documentation Initiative)が採用されていることが少なくない。また、UKDA や FSD 等、多くのアーカイブ機関で、メタデータの記述に必要な専門用語とその定義を含む DDI の Controlled Vocabulary(統制語彙)が採用されており、利用者にとって利便性の高いデータ検索を可能にするデータカタログの開発が進められてきた。

本報告は、このような海外におけるデータのアーカイビングおよびメタデータの整備状況を紹介した上で、日本学術振興会で現在進めているデータカタログの現状と課題について議論したい。

参考文献

UK Data Archive (2007) *Across The Decades: 40 Years of Data Archiving*

¹ 東京大学大学院情報学環(社会科学研究所兼任)

² 中央大学経済学部