

## DDIR を利用した効率的な社会調査 Data Life Cycle

○関西学院大学

中野康人

### 1. 報告の目的

本報告の目的は、社会調査および社会調査にもとづく研究を、より効率的かつ体系的に実践するための「道具」を提供することにある。昨年度の報告（中野 2016）に引き続き、社会調査データの国際規格である DDI(Data Documentation Initiative) とデータ分析環境 R を利用して、効率的に Data Life Cycle を回していく方法を提案する。

### 2. Data Life Cycle と DDI

社会調査の実践は、複数の段階を経ておこなわれる。企画・準備段階、実査段階、分析段階、公表段階など、社会調査全体の過程をここでは Data Life Cycle と呼ぼう。この Data Life Cycle のそれぞれの段階において、われわれは自らが慣れ親しんだ異なるソフトウェアやデータ形式を使って作業を行うことが多い。

しかし、質問項目や回答選択肢など、同じような情報をそれぞれの段階で異なるツールに別々に投入するという作業は、煩雑で非効率的である。多数が関与する大型プロジェクトでは、なおさらその傾向は顕著になる。さらに、調査後に報告書やデータを公開する際に、調査企画段階で参照した先行研究や先行調査の情報が不透明になってしまうことも懸念される。一個人が一貫したツールを使えば、内的な整合性は取れるものの、調査データの効率的な共有や研究成果の reproducibility を考慮すれば、社会調査データに関する「規格」が浸透することが望ましい。

DDI は DDI Alliance (<http://www.ddialliance.org/>) によって提唱されている、社会調査の国際規格である。Data Life Cycle にかかわる情報を、規格化された XML ファイルとして保存し、効率的・統一的にそれらを管理・利用する枠組みである。もともとは、ICPSR などの社会調査データ・アーカイブが、データ収集をする際の規格として整備されたものであるが、社会調査を実践するものにとって、日々の調査・研究活動を効率化する道具となる。

### 3. DDIR とその利用

報告者は、データ分析環境 R において、DDI の規格に基づいた調査データを取り扱う package DDIR を整備してきた。DDIR は、社会調査を「再現可能な研究」Reproducible Research とする道具となり、社会学における知の蓄積に貢献するものである。DDIR を利用すれば、データアーカイブなどから取得した既存の DDI メタデータを読み込んで、分析時に必要な変数名や質問文などの情報を直接参照できる。それに加えて、新バージョンの DDIR では、自らの調査データについて DDI 規格のメタデータを作成することが可能となる。

今回の報告では、Data Life Cycle のそれぞれの場面での DDIR の具体的利用例として、調査票データの準備や取り込み、単純集計レベルの調査報告書自動生成などを紹介する。

#### 【文献】

中野康人, 2016, 「DDI と R を利用した社会調査の統合環境」, 第 89 回日本社会学会大会 (九州大学) .