

エジンバラ大学の研究データ管理について

エジンバラ大学のデータリポジトリ DataShare は 2011 年公開。2012 年 8 月から 2015 年 5 月にかけて「RDM ロードマップ」と呼ばれるプロジェクトに取り組んでいる。このプロジェクトによって RDM サービスを完成させる予定。予算は 130 万ポンド（約 2 億 2,400 万円）。RDM 支援は、研究のライフサイクル全般、(1)研究前の資金獲得(2)研究中のアクティブデータの一時保存(3)研究後のデータ公開(4)非公開データの長期保存(5)データ目録の作成による業績報告まで支援を行っている。さらに、(6)研究者のニーズを把握して(2)から(5)のシステム設計に反映。

研究中に使用するアクティブデータの保存や共有のためのインフラとして、ownCloud を導入。研究者が好む Dropbox とよく似たオープンソース。スタッフや院生を含む全ての研究者に 500GB の領域を無償提供。追加領域を購入することもできる（年間£200/TB）

DataShare、2014 年 3 月の時点 125 アイテム、1,946 ファイル（平均 16 ファイル／アイテム）、合計 76GB のデータを公開データ形式は一般的なもので、plain text（345 ファイル）、x-wav（316）、xml（214）、zip（190）、jpeg（180）と続く。1 ファイルあたりのサイズは 0.1MB 以下のものを除くと 100MB 程度のものが最も多く、最大でも 2GB 程度。研究中にサイズの大きいデータを扱う場合でも、DataShare に登録するのは最終版のみ。サイズが大きい場合は圧縮。さらに大きいデータは研究者自身が学会等に預けているので容量の問題は今のところない。

DataShare のメタデータはシンプル。必須項目はタイトル（title）、登録者（contributor（depositor））、助成機関（contributor.other（funder））、公開者（publisher）、形式（type）、日付（date）。オプションとして作成者（creator）、データ内容（description.table of contents）、言語（language）、権利（rights）、分野（subject）など。図書館員が付与するメタデータは基本的な要素のみで、さらに研究者（研究データの作成者）がデータ記述（data description）を充実させるのが基本。DataCite から取得した DOI を付与する機能があり、各アイテム、もしくはファイルごとに付与。このほか、クリエイティブ・コモンズやオリジナルのライセンスなども表示される

エジンバラ大学の RDM ポリシー、RCUK の OA ポリシーはデータ目録の共有を求めている。そこで 2014 年 11 月には、データ目録のシステムである Data Asset Register を稼働。項目はデータセットのほか、出版物、助成金、研究設備も含まれている。公開用インターフェイスの Edinburgh Research Explorer（<http://www.research.ed.ac.uk/>）は個人単位で表示され、日本の researchmap のようなイメージ。さらに一覧表示される業績目録から、出版物やデータセットへ直接リンク。アドボカシーにも力が入っている。研究者とワークショップや面談を行うことによって、ニーズを把握している。現在研究データの長期保存に向けてソフトウェアの保存やマイグレーションについて検討。