

特集／学ぶ図書館員 2018
MOOC で学ぶオープンサイエンス時代の
研究データ管理支援

古川雅子

近年、論文だけではなく、研究データについても公開し、広く容易に再利用可能にすることにより、研究の加速化や不正防止を効果的に実現することが期待されている。

その実現に不可欠とされているのが、研究を遂行する段階からの適切な研究データ管理（Research Data Management：RDM）である。

研究データ管理とは ある研究プロジェクトにおいて使用された、あるいは生成されたデータの組織化、構造化、保存、共有、公開、再利用に関する一連の作業を指す。

この背景にはオープンサイエンスの潮流がある。2013年のG8科学大臣会合における研究データのオープン化に関する共同声明を皮切りとして、国内でもオープンサイエンスに関する議論が活発化している。2018年6月29日には、内閣府の「国際的動向を踏まえたオープンサイエンスの推進に関する検討会」から、「国立研究開発法人におけるデータポリシー策定のためのガイドライン」が公開された。

こうしたデータポリシーの策定は、研究機関だけではなく、大学にも展開されることが予想される。これまで機関リポジトリを運用してきた大学図書館には、これからの時代に必要となる研究データ管理の支援についても大きな期待が寄せられている。しかしながら、このような支援サービスを実現するためには、具体的に何をどのようにサポートすべきか知らなくてはならない。

たとえば、海外では、既に研究データに関してさまざまなオンライン教材が公開されており、エジンバラ大学MANTRAは、研究プロジェクトの一環としてデジタルデータを管理する人を対象とした無料のオンラインコースを提供している。

FOSTERでは、オープンサイエンスの詳細を知るために必要なeラーニングコースが集められており、Figshareは、研究データを公開するためのレポジトリであるとともに、研究データ管理に関する教材も公開している。

日本では、オープンアクセスリポジトリ推進協会（JPCOAR）の研究データタスクフォースが海外の動向も

調査しながら研究データタスクフォースが海外の動向も調査しながら研究データのライフサイクル（生成、加工、分析、保存、公開、再利用）など研究データ管理の基礎を学ぶ教材の開発を行い、2017年6月6日付でJPCOARの公式ウェブサイト上でCC-BY（クリエイティブ・コモンズ・ライセンス）で公開した。各章は、スライドと解説と確認テストで以下のように構成されている。

第1章：導入編であり、RDMが必要とされる背景や、研究データおよびRDMの定義について解説している。

第2章：DMP（Data Management Plan）の定義や動向について述べた上で、実際のDMPの策定方法を解説する。あわせてDMPを支援するツールについても紹介している。

第3章：研究データの保存と共有をテーマとし、長期保存にあたってのセキュリティ上の留意事項や、共有・再利用のためのデータリポジトリの活用などについて解説している。

第4章：研究データの組織化、文書化、メタデータ作成をテーマとし、主要なメタデータスキーマを紹介するなど、データを再利用に供するためのシステムティックな管理方法を解説している。

第5章：法・倫理的問題をテーマとし、著作権に関するライセンスや、研究不正に関するポリシーについて解説している。

第6章：研究者が研究を進めるにあたってのポリシーをテーマとしており、研究不正に関するポリシーの事例と、そこでの研究データ取扱いについて解説している。

第7章：RDMサービスの設計をテーマとし、RDMを支援する人材の確保や、研究データ保存・公開の基盤システムについて解説している。

教材の公開後は、さらなる有効活用と普及を目指し、国立情報学研究所と共同でJMOOCプラットフォームのgacco上で、RDMトレーニングツールを活用した講座を制作した。

このMOOC講座「オープンサイエンス時代の研究データ

管理」は、2017年11月15日から2018年1月15日まで開講された。MOOCでは、RDMトレーニングツールで7章の構成だった教材を4週間分のコースに再編成し、研究データタスクフォースの協力を得て映像も補完した。各週は、4~5本の映像、ダウンロード教材、確認テスト(10問の選択問題)で構成した。この他に、任意で回答する開始前/実施後アンケートおよび、ディスカッションボードを設置した。また、各週の確認テストの合格点が7割に達した場合には修了証を発行した。

開発した研究データ管理オンラインコースの受講者数は2,305名、修了率は25%であった。gaccoの平均修了率は15%であり、一般的にMOOCの修了率は10%前後という中では高いと言えるだろう。

確認テストの得点分布を見ると、登録のみで結局受講しないことによる0点を除けば、70%が修了基準であるのにも関わらず合格100点で修了した受講者が最も多かった。

開始アンケートの回答は必須ではないものの、770名から回答を得た。回答者の6割程度が男性であり、年齢層はほぼJMOC全体との構成と同じだった。勤務形態はフルタイムと回答した者が616名であり、勤務先は大学研究機関が50%を占めた。その中でも特に多かったのは図書系と回答した者で、回答者全体の30%を占めていた。

受講後アンケートでは、「講座の内容は、あなた自身にとって有用でしたか」という質問に対して、「大変有用である」、「有用である」の割合が9割程度だったが、図書系とそれ以外の者を比較すると、図書系の受講者は、「大変有用である」と答えた割合が更に多かった。

また、そのように回答した理由を聞いたところ、図書系の受講者の回答を見ると、「オープンサイエンスの基本的な考え方について学ぶことができた」、「図書館の現状について認識できた」等の記述がある反面、「図書館と他の部局との連携が課題」といった今後の実施に関して具体的な検討の記述も見られた。なお、学習ログの分析結果からは、図書系の受講者は、それ以外の者に比べて熱心に映像の視聴や確認テストの回答を行う、成績の良い受講者の割合が多いことが示され、図書系学習者の真面目さと研究データ管理支援への関心の高さが伺えた。

現在、オンライン講座第2弾として、研究支援者、研究者を支援する立場の職員、基盤センター技術系スタッフなどの支援者向けに、研究プロセス(研究前、研究中、研究後)に沿ってどのようなサービスをしたらいいのか、そのデザインはどうしたらいいのか、実践するにはどうしたらいいのかな

どを学べる教材を準備中である。時代の要請に応えるべく自己研修を行う際には、ぜひこのようなオンライン教材を活用していただければと思う。

また、個人の学びに留まらず、ぜひ図書館でリテラシー教育のプログラムの中に研究データ管理を組み込んで、特に大学院生や若手の研究者向けに研究データ管理について学んでもらえるような場を作ったり、学内研修として研究者、URA、技術系職員等を対象とした研修会のようなものを開いて意識を共有できるような場を作ったりと、実際に図書館業務の発展に活かしていただければと思う。

参考文献

- 1) 船守美穂. 2017. オープンサイエンス推進に関わる学術機関の役割と課題. 情報知識学会誌, vol.27, no.4, pp.309-322. DOI: https://doi.org/10.2964/jsik_2017_034
- 2) 国立研究開発法人におけるデータポリシー策定のためのガイドライン (Published June 29, 2018 by Cabinet Office, Government of Japan), <http://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.pdf>
- 3) MANTRA, <https://mantra.edina.ac.uk/>
- 4) Orth, A., Pontika, N., Ball, D. 2016. FOSTER's Open Science Training Tools and Best Practices. IOS Press. DOI:<https://doi.org/10.3233/978-1-61499-649-1-135>
- 5) Figshare, <https://knowledge.figshare.com/open-data/about-rdm>
- 6) 常川真央, 天野絵里子, 大園隼彦, 西園由依, 前田翔太, 松本侑子, 南山泰之, 三角太郎, 青木学聡, 尾城孝一, 山地一禎. 2017. 研究データ管理(RDM)トレーニングツールの構築と展開. 情報知識学会誌, vol.27, no.4, pp.362-365. DOI: https://doi.org/10.2964/jsik_2017_042
- 7) JMOC, <https://www.jmoooc.jp/>
- 8) 古川雅子, 尾城孝一, 山地一禎. 「研究データ管理オンライン講座の開発と受講者特性の分析」, 情報処理学会情報教育シンポジウム2018論文集. 2018.