

東邦大学学術リポジトリ

Toho University Academic Repository

タイトル	「オープンサイエンス時代の研究データ管理」受講報告
別タイトル	Report on Taking a Course Entitled "Research Data Management in Open Science"
作成者(著者)	亀井, 雄大
公開者	日本医学図書館協会
発行日	2018.09.01
ISSN	04452429
掲載情報	医学図書館. 65(3). p.154-157.
資料種別	学術雑誌論文
内容記述	事例報告
著者版フラグ	publisher
メタデータのURL	https://mylibrary.toho-u.ac.jp/webopac/TD17558347

「オープンサイエンス時代の研究データ管理」受講報告

亀井 雄大*

東邦大学医学メディアセンター

I. はじめに

国際的なオープンサイエンスの潮流の中で、論文の根拠となる研究データをいかに管理するかが課題となっている。日本では大学や研究機関における組織的な研究データ管理は未だ行われておらず、この実現のためには研究データ管理を担う人材の育成が必要な状況である。

このような背景をもとに開講されたのが、本稿で取り上げるオンライン講座「オープンサイエンス時代の研究データ管理」¹⁾である。本講座は受講者がオープンサイエンスで求められる「研究データ管理 (Research Data Management: RDM)」の基礎的な知識を習得し、所属機関において研究データ管理をサービスとして提供できるようにすることを目的に開講された。

本稿では、筆者が本講座を受講して分かったこと、感じたことの中からポイントを絞って報告する。なお本稿の内容は、2018年6月に開催された日本医学図書館協会総会の分科会A「オープンサイエンス」で発表した内容を再構成したものである。

II. 受講の目的

筆者は2016年に東邦大学医学メディアセンターに入職し、資料の受入や目録作成、機関リポジトリへの大学紀要の登録を行っている。2016年に文部科学省²⁾は、研究データ管理に機関リポジトリを活用することや研究データ管理の支援者としての大学図書館員の役割について言及している。このような事情から、大学図書館員として今後、研究データ管理の支援に携わる可能性が高いと考え、オープンサイエンスと研究データ管理に関する基礎知識の習得を目的として本講座を受講した。また次の3点について学習できるのではないかと期待した。

①オープンサイエンスの中で大学図書館員に期待される

役割について、②機関リポジトリにおける研究データの取り扱い方について、③医学図書館員に求められる研究データ管理の知識について。

III. 講座について

本講座は日本におけるMOOC (Massive Open Online Courses: 大規模公開オンライン講座) のプラットフォームの1つである「gacco」で2017年11月15日から2018年1月15日まで開講された³⁾。大学図書館員やURA (University Research Administrator: リサーチ・アドミニストレーター)、データキュレーターなどの研究支援者が受講者として想定されていた。授業はweek1からweek4までの4週構成となっていた。開講時にweek1が公開され、それ以降は毎週水曜日に次の授業が公開された。それぞれの授業の勉強時間は4~5時間程度とされていた。また授業内容に沿った課題が用意されており、授業の公開後2週間以内に提出することが修了の条件となっていた。授業はパワーポイントで作成された教材と動画コンテンツであり、受講者はこれらをダウンロードすることも可能であった。

2018年7月現在、本講座はgacco上では利用できなくなっているが、当時公開された教材と動画コンテンツが国立情報学研究所 (以下、NII) のウェブサイトを利用可能である¹⁾。

IV. 分かったこと

1. 研究データ管理とは

研究データ管理には次の3つのポイントがある。

1) 研究前におけるデータ管理計画の策定、2) 研究中におけるデータの取り扱い、3) 研究後におけるデータの取り扱い。

1) 研究前におけるデータ管理計画の策定

研究データの適切な管理にはデータ管理計画 (Data Management Plan: 以下、DMP) が不可欠である。本講座ではDMPを「研究の実施段階から終了後に至るまで

*Yudai KAMEI: 〒143-8540 東京都大田区大森西5-21-16.
yuudai.kamei@mnc.toho-u.ac.jp (2018年7月21日 受理)

の期間において、研究データがどのように生成、管理、保存される予定かを文書化したもの」としている。研究者はDMPを策定し、それに従った研究データの管理をすることで、自身の研究の質や効率性、信頼性を高めることができる。また、研究中に研究データが損失するリスクを軽減することができる。さらに研究データを利活用可能な状態で長期的に保存することができる。

研究助成機関が助成の要件として、研究者にDMPの提出を義務付ける動きが国際的に高まっている。日本でも科学技術振興機構（JST）や日本医療研究開発機構（AMED）がDMPの提出を義務化している⁴⁾⁵⁾。各助成機関が提出を求めるDMPはそれぞれ個別の様式をもっており標準化されていないが、各助成機関の要件に則った計画書を作成する支援ツールとして、カリフォルニア大学が提供するDMPTool⁶⁾やDCC（Digital Curation Center）が提供するDMPOnline⁷⁾などが存在している。

2) 研究中におけるデータの取り扱い

研究者は日々生成する研究データを適切に管理するために、実験ノート等の形で研究データに関する文書をまとめる必要がある。これは自分自身の記録のためであると同時に、第三者の再利用をより容易にするためである。またそれは自らの研究活動が公正であることを証明するものとなる。記録する内容は、データの収集方法や処理方法などである。

生成した研究データは安全に保管する必要がある。そのために配慮すべきこととしては、データの喪失を防ぐための定期的なバックアップや、流出や紛失を防ぐためのセキュリティ対策といったものが挙げられる。

研究の進展につれて研究データが多くなると、管理や識別が困難になる。一貫したファイルの命名ルールを決めておくなどの方法で研究データを体系立てて管理していくことが重要である。

3) 研究後におけるデータの取り扱い

研究データの保存期間は日本学術会議によると、原則、論文の発表後10年間とされている⁸⁾。これは主に研究の不正対策を目的に設けられた期間である。研究前に策定するDMPにおいても、データセットの長期保存計画を定めておく必要がある。また研究データを将来にわたって利用可能とするために、特定のソフトウェアに依存していない、広く普及しているファイルフォーマットを選ぶことが重要である。

研究データの公開は、研究の公正さを社会に示すことにつながる。このため生成した研究データのうち、特にコアとなるエビデンスデータを中心にオンラインでの公

開を義務付けている機関もある。

公開される研究データには、検索性を向上させるためにメタデータを付与する。研究データに付与するメタデータ・スキーマの例としてDataCiteとDublinCoreが定めたものがある。メタデータにコンテンツの識別子であるDOIや研究者の識別子であるORCIDなどを付与することで、研究データの検索性が向上し、アクセスされやすくなる。また公開した研究データには、ほかの研究者の再利用が容易になるように、クリエイティブ・コモンズ等が提唱しているライセンスルールを用いて、コンテンツの利用許諾の条件を明示することも重要である。

しかし、研究データの中には公開に適さないものもある。非公開を検討すべき研究データとして、例えばプライバシー情報などを含むものがある。研究データのもつ諸側面に配慮し、公開や制限付き公開、非公開を決めることを、「オープン・クローズ戦略」と呼ぶ。機関や研究者は、研究データを公開する際には単にオープン化を推進するだけでなく、この戦略について十分に検討することが重要である。

2. 研究データ管理の意義について

公開された研究データは再利用できることが重要である。そのためには、研究データを第三者が検索可能かつ内容を解釈可能な状態に整備する必要がある。適切に研究データ管理を行うことでこれを実現することができる。

研究者は、研究データ管理を通して研究の質や効率性を高めることができる。また適切な研究データ管理は研究の正確性や信頼性を保証することにもつながる。公開することを想定した研究データ管理は、研究の不正対策に効果を発揮する。

公開された利活用可能な研究データが増加することで、他の研究者や企業が同じようなデータを再度収集する必要がなくなり、研究資金や資源、労力の節約につながる。また研究データの再利用者が新たな観点から別の研究成果を生み出すことも可能となる。

3. 研究データ管理の環境構築について

研究データに関するポリシーを定めている国内の大学機関に京都大学⁹⁾や大阪大学¹⁰⁾、九州大学¹¹⁾がある。これらのポリシーには、研究データの取り扱いの方針として「長期保存や公開の対象となる研究データの種類」、「データの保存期間」、「研究データの帰属先」などが定められている。

研究データに関わる支援体制を整えるためには、所属

機関内における各ステークホルダーの状況や、機関内の研究活動、研究者のニーズを把握することが重要である。

研究データ管理の支援サービスとして想定されるものに、DMPの策定支援や、機関リポジトリを活用した研究データの保存や公開、研究データを公開する際のメタデータの作成支援などがある。

V. 分からなかったこと

受講を通して研究データ管理の基礎的な知識を学習することができた。一方で受講前に期待していた3点については分からないことも多かった。そもそも本講座は、大学図書館員のみを対象としたものではなかったため、職種を限定した実務的な情報は多くなかった。

1. オープンサイエンスの中で大学図書館員に期待される役割について

本講座では、各研究支援者の職業種別ごとの役割については触れられなかった。そのため研究データ管理において、大学図書館員がどこまでの領域をカバーすべきか判然としなかった。大学図書館員に期待されている役割を把握することで、より焦点を絞った学習が可能となると考える。

2. 機関リポジトリにおける研究データの取り扱い方について

機関リポジトリで研究データを扱うための知識が得られるのではないかと期待していたが、機関リポジトリで研究データを扱うことの重要性について分かった一方で、実務に関係するような具体的なことは言及されなかった。機関リポジトリで研究データを扱えるようになるためには、準備からサービス提供までに相応の時間を要すると考える。事前にできる限り準備しておくためにも、機関リポジトリにおける研究データ管理の方法について解説が欲しかった。

3. 医学図書館員に求められる研究データ管理の知識について

本講座では、研究分野別の課題については触れられなかった。しかし医学図書館員が扱う研究データのほとんどは、医学系の研究分野から生成されたものになると考える。そのため医学図書館員としては、特に医学系の研究データを扱う際の注意点を把握しておきたかった。

本講座の中で、個人情報を含む研究データは取り扱いを注意しなければならないデータの1つとして紹介され

ていた。医学分野の研究データには被験者のプライバシー情報が含まれる可能性が高いため、この点は医学図書館員が研究データを取り扱うときの注意点の1つであると考えられる。医学図書館員が安心して研究データを取り扱うことができるように、分野特有の課題について議論し知識を深めていく必要があると感じた。

4. 国内の動向について

参考として紹介されているのは主に海外の先進的な事例であった。現在、国内ではどのような議論や活動が行われているのか詳しくは分からなかった。例えば、本講座では触れられていなかったが、現在NIIの組織下にオープンサイエンス基盤研究センターが設立されており、研究データの公開や管理、検索のための新しい基盤の開発を行っているとのことである¹²⁾。こういった国内の機関の取り組みについても紹介してほしい。

VI. 受講した感想

内容は体系的にまとまっており分かりやすかった。また動画を見ながら学習を進められるため、学習時間の調整もしやすいと感じた。さらに教材や動画はすべてダウンロード可能になっており、職場や自宅などでパソコンを使って受講するだけでなく、移動中にタブレットやスマートフォンで学習することも可能であり便利だと感じた。

gaccoでの公開中はディスカッションの場が用意されており、図書館員だけでなく研究者や企業関係の方など様々な立場の参加者が意見交換を行っており刺激的であった。講師や運営側の有識者が質問に答えていることもあり、参考にもなっていたので閉講と同時に閲覧できなくなってしまったのは残念である。

また設定されていた開講時間が短く感じた。本講座の修了証をもらうためには毎週4時間程度受講のための時間を割く必要があり、かつ内容に沿った課題を各授業の公開後2週間以内に提出しなければならなかった。

VII. おわりに

本講座はオープンサイエンスや研究データ管理について基礎的な知識を習得したいと考えている研究支援者に推薦できる内容であった。ただし大学図書館員や医学図書館員の役割についての詳細な言及はなかった。また研究データ管理に関する医学分野特有の課題についても述べられてはいなかった。このため本講座から得られる知識のほかにも自ら情報を収集していくべきだと感じた。本講座で得られた知識を土台とし、更なる学習を進めて

医学図書館における研究データ管理の支援のあり方について検討していきたい。

引用・参考文献

- 1) 国立情報学研究所. オープンサイエンス時代の研究データ管理 [internet]. <https://www.nii.ac.jp/service/jmooc/rdm/> [accessed 2018-07-08]
- 2) 文部科学省. 学術情報のオープン化の推進について (審議まとめ) その1 [internet]. http://www.mext.go.jp/component/b_menu/shingi/toushin/_icsFiles/afieldfile/2016/04/08/1368804_1_1_1.pdf [accessed 2018-07-19]
- 3) gacco. オープンサイエンス時代の研究データ管理 [internet]. <https://lms.gacco.org/> [accessed 2018-07-08]
- 4) 科学技術振興機構. JSTからのお知らせ オープンサイエンス促進に向けた研究成果の取扱いに関するJSTの基本方針. 情報管理. 2017;60(2):143-5.
- 5) 日本医療研究開発機構. データマネジメントプランの提出の義務化について [internet]. <https://www.amed.go.jp/koubo/datamanagement.html> [accessed 2018-07-06]
- 6) University of California Curation Center of the California Digital Library. DMPTool [internet]. <https://dmptool.org/> [accessed 2018-07-08]
- 7) Digital Curation Center. DMPonline [internet]. <https://dmponline.dcc.ac.uk/> [accessed 2018-07-08]
- 8) 日本学術会議. 科学研究における健全性の向上について [internet]. <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-k150306.pdf> [accessed 2018-07-06]
- 9) 京都大学. 京都大学における公正な研究活動の推進等に関する規程第7条第2項の研究データの保存, 開示等について定める件 [internet]. http://www.kyoto-u.ac.jp/ja/research/ethic/research_guide/documents/research_data150730.pdf [accessed 2018-07-06]
- 10) 大阪大学. 大阪大学における研究データの保存等に関するガイドライン [internet]. http://www.osaka-u.ac.jp/ja/research/researchintegrity/files/data_guideline.pdf [accessed 2018-07-06]
- 11) 九州大学. 研究データの保存等に関するガイドライン [internet]. <https://www.kyushu-u.ac.jp/f/1461/guideline.pdf> [accessed 2018-07-06]
- 12) 国立情報学研究所オープンサイエンス基盤研究センター. サービス概要 [internet]. <https://rcos.nii.ac.jp/service/> [accessed 2018-07-08]

Report on Taking a Course Entitled “Research Data Management in Open Science”

Yudai KAMEI

Toho University Medical Media Center. 5-21-16, Omori-Nishi Ota-ku, Tokyo 143-8540, Japan

Abstract: As part of the movement towards open science, academic librarians including medical librarians are expected to engage in Research Data Management (RDM). The MOOC (Massive Open Online Course) “Research Data Management in Open Science” course discussed in this report was created by the NII (National Institute of Informatics) in November 2017 to enable research supporters to acquire basic knowledge regarding RDM. Students are able to learn the basic principles of RDM

through this course. On the other hand, information regarding domestic trends and subjects in specific fields, such as life science, are not included in the course. Medical librarians must acquire expertise regarding research data in the medical field in addition to the knowledge gained from this course.

Keywords: Open Science, Research Data Management, Data Management Plan
(*Igaku Toshokan*. 2018;65(3):154-157)