

研究データのオープン化と そのメリット

2020(令和2)年4月

国立大学図書館協会オープンアクセス委員会



この資料はクリエイティブ・コモンズ 表示 4.0 国際ライセンスの下に提供されています
<https://creativecommons.org/licenses/by/4.0/>

研究データとオープン化の流れ

研究データ

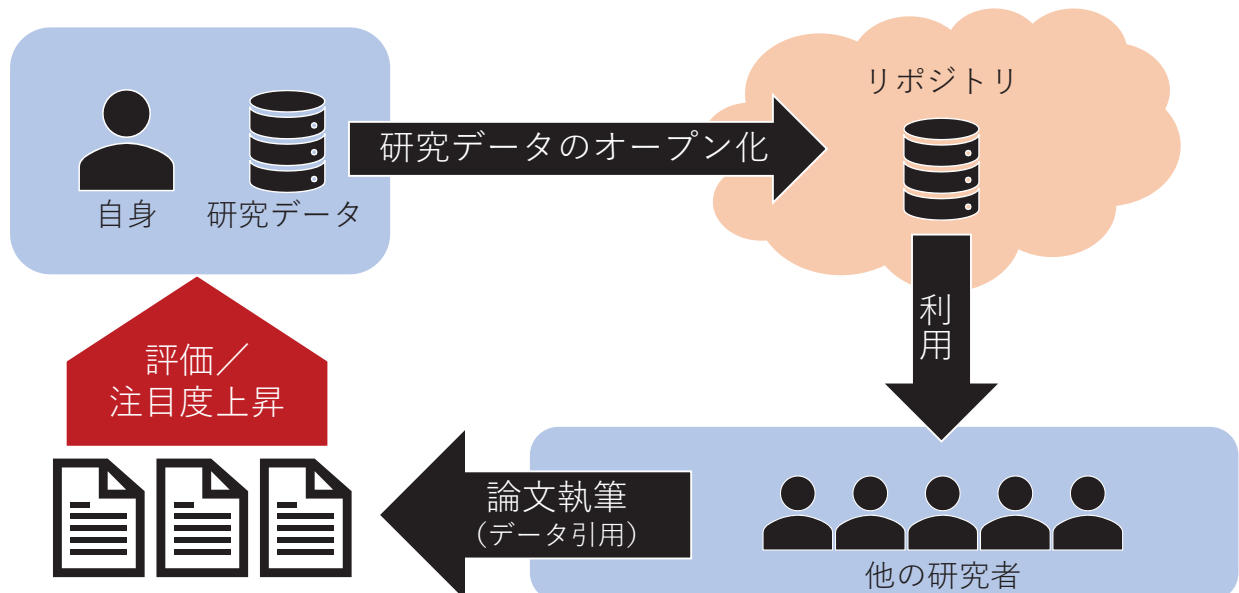


- 研究の過程で、あるいは研究の結果として収集・生成される情報
- 形態は、数値、画像、テキスト、コンピュータプログラムなど様々

- 研究活動の第一の成果は「学術論文」(* 学術図書等も含む)だが、近年「研究データ」も**研究成果として注目**されている
- 新たな研究活動の促進のため、国内外で「研究データ」の**リポジトリ*注1)等でのオープン化(インターネット等による公開)の流れが加速**している
 - 2013年6月、英国で開催されたG8科学大臣会合の共同声明において、研究データのオープン化が言及され、世界的に議論が加速する契機に*注2)
 - 我が国でも2015年3月、内閣府の検討会が報告書「我が国におけるオープンサイエンス推進のあり方について」*注3)を公表し、研究データのオープン化に関する検討が進みつつある

オープン化のメリット

研究データのオープン化により、他の研究者による利用（データ引用）や新たな研究成果が生まれ、結果として、**自身の研究に対する評価や注目度の上昇につながる**



3

研究データの具体例及びオープン化の状況

オープン化が既に一般的となっている研究データもあれば、そうでないものもある

オープン化済

- 物質の結晶構造データ *注4)
- 遺伝子等の塩基配列データ *注5)
- 自然現象の観測データ（宇宙、海洋 *注6)、気象、地震等)
- 統計調査・社会調査等の個票データ *注7)
- 古典籍の全ページ画像 *注8)

特定の分野や専門機関等において、公開によるメリットが存在するとの共通認識が存在

一部オープン化済

- 論文の根拠データ

オープン化を「必須」or「推奨」とするジャーナル *注9) や研究助成機関 *注10) も存在

オープン化未開拓

- 使用済データ (*論文の根拠データを除く)
- 未使用データ (うまくいかなかった研究のデータを含む)

4

懸念とその対応①

どのようなデータを公開すればよいのか

他の研究者による再利用が期待できるもので、公開のために特別な手続き・検討等が不要な、容易に公開可能なものから公開するとよい

公開ができない／困難なデータはどんなものか

共同研究者の合意、企業等との共同研究契約、法的問題（個人情報、特許等の知財、安全保障等）、研究コミュニティの慣習（例えば、治験データの扱い、絶滅危惧種の生息地域、文学研究のインタビュー等）などの点で問題があるデータ *注11)

どこで公開すればよいのか

機関リポジトリや【※各機関で実態に合わせて記述】、ジャーナルや研究助成機関が推奨するリポジトリ。また、データリポジトリの検索サイト re3data.org (<https://re3data.org/>) で適切なりポジトリを探すこともできる。適当なりポジトリが見つからない場合は、figshare、Zenodo等の汎用リポジトリが利用できる

5

懸念とその対応②

どのくらいのサイズの研究データに対応できるのか

リポジトリ毎に異なるため、各リポジトリの管理者に問い合わせるか、FAQ等で上限に関する記述を確認するとよい

公開には手間がかかる

機関リポジトリで公開する場合は、担当部署（図書館等）がサポートする
【※各機関で実態に合わせて記述】

公開したデータを他の研究者が利用するのは難しいのでは？

利用しやすくするため、データの内容を説明するメタデータ等を付与する必要がある。機関リポジトリで公開する場合は、担当部署（図書館等）がメタデータの作成等をサポートする 【※各機関で実態に合わせて記述】

研究データをどのように引用するのか

各リポジトリでは、DOI等の永続的なURLが付与されるため、引用しやすくなっている *注12)。また、研究データの引用ルールなどの検討が進んでいる *注13)

6

- 1) 「リポジトリには、大学・研究機関の成果を保存・管理・公開することを主眼とする『機関リポジトリ』、分野における研究資源を保存して活動促進に資する『分野リポジトリ』、分野や機関を限定しない『汎用リポジトリ』などの分類がある」(p.2, 脚注1)
 【文献】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(内閣府). 研究データリポジトリ整備・運用ガイドライン. 2019 (<https://www8.cao.go.jp/cstp/tyousakai/kokusaioopen/guideline.pdf>)
- 2) 以下文献p.3
 【文献】国際的動向を踏まえたオープンサイエンスに関する検討会. 我が国におけるオープンサイエンス推進のあり方について:サイエンスの新たな飛躍の時代の幕開け. 2015 (<https://www8.cao.go.jp/cstp/sonota/openscience/>)
- 3) 上記2)の文献
- 4) X線回折によって得られる物質の結晶構造データは、「単にその物質の結晶構造を明らかにするだけでなく、その物質が有する物性・特性の起源を演繹・帰納するのに欠かせない情報」(p.178)であり、また「より高い物性・特性値を示す物質の開発において、類縁もしくは他の物質の情報を知る事も非常に重要な知見」(p.178)である。そのため「既知の結晶構造データの共有化も早くから図られて」(p.178)きた。
 【文献】松下能孝. 結晶構造データベースと結晶学共通データ・フォーマットCIFについて: 1. 結晶構造データベース. Journal of Surface Analysis. 2013, 19(3), p. 177-187 (https://www.jstage.jst.go.jp/article/jsa/19/3/19_177/article/-char/ja/)

- 5) 「ヒトをはじめとする多くの生物種の遺伝子やゲノムの塩基配列が急激な勢いで決定され、国際塩基配列データベース共同体(International Nucleotide Sequence Database Collaboration; INSDC, <http://www.insdc.org>)に登録されている」「世界中の人々は誰でも公開されたデータを自由に利用でき、それを使って研究、教育や事業などを進めることができる」(p.182)
 【文献】舘野義男. ライフサイエンス分野の統合データベース(第7回)海外データベースとの連携: International collaboration among DDBJ, EMBL Bank and GenBank. 蛋白質核酸酵素. 2008, 53(2), p.182-189 (<http://ci.nii.ac.jp/naid/40015830835/ja/>)
- 6) 「海洋観測データ(主として海洋物理・化学、地球物理系)では50年以上前からデータ共有の仕組みが存在」(p.3)
 【文献】JAMSTECにおける海洋データ管理について(内閣府 オープンサイエンス推進に関するフォローアップ検討会(第5回)配布資料資料2)(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(第1回)参考資料3)
 (<https://www8.cao.go.jp/cstp/tyousakai/kokusaioopen/1kai/sanko3.pdf>)
- 7) 我が国では以下のような機関で収集・提供している。
 - ・ 東京大学社会科学研究所 附属社会調査・データアーカイブ研究センター
 - ・ JILPTデータ・アーカイブ(独立行政法人 労働政策研究・研修機構)
 - ・ 慶応義塾大学パネルデータ設計・解析センター

- 8) 例えば「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」（国文学研究資料館）で作成・提供されている。この計画は、「国内外の大学等と連携して、古典籍約30万点の全冊画像化を行い、当館が構築してきた古典籍の書誌データベースと統合して、自在に画像を検索できる「新日本古典籍総合データベース」という研究基盤を作り、その画像を用いて国際的な共同研究のネットワークを構築する」ものである。
【文献】日本語の歴史的典籍の国際共同研究ネットワーク構築計画（略称：歴史的典籍NW事業）パンフレット（令和元年8月改訂版）
(https://www.nijl.ac.jp/pages/cijproject/images/NIJL-NWproject_pamphlet.pdf)
- 9) 2014年、インパクトファクターの高い学術雑誌220誌について、研究データの共有（オープン化）ポリシーを調査した結果、「必須」が100誌（45.5%）、「推奨」が30誌（13.6%）であった（以下文献p.25）
【文献】池内有為, 逸村裕. 分野別データ共有ポリシーと雑誌の特徴：分野間比較と特徴分析. 日本図書館情報学会誌. 2016, vol. 62, no. 1, p. 20-37
(https://doi.org/10.20651/jslis.62.1_20)
- 10) 科学技術振興機構（JST）は、エビデンスデータの公開を推奨している（以下文献p.3）
【文献】科学技術振興機構. 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. 2017 (<https://www.jst.go.jp/all/about/houshin.html#houshin04>)
- 11) 以下文献p.8～10が参考になる
【文献】研究データ利活用協議会研究データライセンス小委員会. 研究データの公開・利用条件指定ガイドライン. 2019
(https://japanlinkcenter.org/rduf/doc/rduf_license_guideline.pdf)

- 12) 研究データへのDOI付与については、以下の文献が参考になる。なお、GenBank、CSD（Cambridge Structural Database）などの分野リポジトリでは、DOI等ではなく、固有の番号等が使用されている
【文献】研究データ利活用協議会. 研究データにDOIを付与するには？：5分で分かる研究データDOI付与. 2019 (https://japanlinkcenter.org/rduf/doc/rduf_rdc_doileaflet.pdf)
- 13) 例えば以下文献のような宣言が出されており、120以上の出版社（Elsevier、Nature Publishing、Wiley等）及び学協会等が支持を表明している
(<https://www.force11.org/datacitation/endorsements>)（2020-03-17参照）
【文献】FORCE11: Data Citation Synthesis Group (2014). Joint Declaration of Data Citation Principles. データ引用原則の共同宣言：最終版. (研究データ利活用協議会: リサーチデータサイテーション小委員会, Trans.)
(https://japanlinkcenter.org/rduf/doc/rduf_rdc_iddcp_ja.pdf)