

## CA2025

## 再現性・複製可能性と研究図書館

にしおか ちふみ\*  
西岡千文\*

## 1. はじめに

2000年代後半以降、心理学等様々な分野で、研究の再現性 (reproducibility)・複製可能性 (replicability) の危機に対する問題意識が高まっている<sup>(1)</sup>。2021年に採択されたユネスコのオープンサイエンスに関する勧告<sup>(2)</sup>(E2485参照)では、オープンサイエンスの価値を実現する枠組みを提供する原則の一つとして再現性が挙げられており、再現性は健全な科学の発展に不可欠な要素であるという共通認識が形成されつつある。

本稿では、研究の再現性・複製可能性の定義とともに、再現性・複製可能性において研究図書館が果たす役割を概観する。次章ではまず、再現性・複製可能性の定義を確認する。続いて、再現性・複製可能性の危機を引き起こす要因とそれらに対する取り組みを報告する。4章では、それらを踏まえて、研究図書館が果たす役割について述べる。

## 2. 再現性と複製可能性の定義

再現性や複製可能性という用語をめぐるのは、曖昧で一貫性のない使用が指摘されており、混乱が生じている<sup>(3)</sup>。最もコンセンサスを得られている定義として、全米科学・工学・医学アカデミー (NASEM) の報告書「科学における再現性と複製可能性 (Reproducibility and Replicability in Science)」<sup>(4)</sup>(2019年5月公開)によるものが挙げられる。報告書では、再現性と複製可能性は以下のように定義されている。

- ・ **再現性**とは、同一の入力データ、計算ステップ、方法、コード、分析条件を使用して一貫した結果を得ることを指す。この定義は、計算的な再現性 (computational reproducibility) と同義であるとされる。
- ・ **複製可能性**とは、同一の科学的質問に答えることを目的とした研究で、一貫した結果を得ることを指す。各研究は独立してデータの取得を行う。

研究を再現できないことは研究の厳密さが不足していることを示す。対して、複製可能でないことは、自然界における複雑性等が影響していることもあり、複製の失敗は新しい発見に繋がることもある。

なお、グッドマン (Steven N. Goodman) ら<sup>(5)</sup>の定

義においては、再現性は「方法の再現性」に該当し、複製可能性は「結果の再現性」とされている。よって、再現性は複製可能性の前提条件として捉えることもでき、NASEMの定義における再現性と複製可能性の両方の概念を表すこともある。心理学分野での大規模な複製可能性の検証の試みである Open Science Collaboration による Reproducibility Project<sup>(6)</sup>でも、複製可能性を意図して“reproducibility”が使用されている。本稿では、NASEMの定義に従って、再現性と複製可能性を区別する。

NASEMの報告書に基づき、米国情報標準化機構 (NISO) は、出版関係者が中心となって、「再現性に関するバッジについての推奨指針」<sup>(7)</sup>(2021年1月公開)の策定を行った。出版者は、再現性やオープン性等のステータスを示すために、論文等へバッジの付与を行っている。推奨指針では、バッジスキームの標準化を行っており、以下の4つのバッジが定義されている。

## ・ Open Research Object (ORO)

研究で作成・使用したデータやソースコード等のデジタルオブジェクトが公開リポジトリに永久的にアーカイブされていることを指す。リポジトリが、識別子の割り当て、永続性の保証を行っており、デジタルオブジェクトは、標準のオープンライセンスによって可用性が高い状態であることを示す。

## ・ Research Objects Reviewed (ROR)

著者によって作成・使用された関連する全てのデジタルオブジェクトが、バッジの発行者が定める条件に沿ってレビューされたことを示す。これには、デジタルオブジェクトの動作のみの確認も、再現性の確認 (次の項目) も該当する。

## ・ Results Reproduced (ROR-R)

RORに加えて、独立した主体によって著者が作成した研究オブジェクト、方法、コード、分析の条件を使用して計算結果を再生成できることを証明する手順が、バッジ発行者によって実行されたことを指す。すなわち、再現性が証明されたことを示す。

## ・ Results Replicated (RER)

同一の科学的質問に対する独立した研究が、同一の結果を導く結果を得たことを指す。よって、複製可能であることを示す。

ORO以外は、デジタルオブジェクトがオープンである必要はないとされる。

\* 国立情報学研究所

### 3. 再現性と複製可能性の危機の要因と解決に向けた取り組み

NASEMの報告書<sup>(8)</sup>では、再現性の危機の要因として、データやコードが公開されないこと、計算環境や依存するソフトウェア等の詳細についての記録の不足、デジタルリソースの陳腐化などが挙げられている。

複製可能性の危機の要因として、ロメロ (Felipe Romero)<sup>(9)</sup>は、まず出版バイアス、さらに出版バイアスに関連してp値ハッキング<sup>(10)</sup>等の問題のある研究慣行 (questionable research practice : QRP) 等を挙げている。出版バイアスは、仮説を支持する結果や統計的に優位な結果が得られた場合に偏って研究成果が出版されることを指す。一般的な帰無仮説有意性検定の枠組みでは、誤った仮説は最大5%の確率で有意になり得る。有意な結果のみが公開されるとすると、偽陽性である5%の結果のみが公開されていることになり、真陰性である残りの95%はお蔵入りすることになる。ポパー (Karl Raimund Popper) は『科学的発見の論理』<sup>(11)</sup>で、「再現できぬ1回だけの出来事は、科学にとっていかなる意義もない」と述べているが、出版バイアス下では「再現できぬ1回だけの出来事」が出版され続けることになる。よって、研究を複製することは難しい。その上、出版バイアス下では、QRPのインセンティブが発生する。このことも、複製を難しくする。

さらに、再現性と複製可能性の危機に共通した根本的な要因として、科学の報酬システムが挙げられている<sup>(12)</sup>。科学の報酬システムの根幹には先取ルール (priority rule)<sup>(13)</sup>、すなわち発見を行った最初の研究者のみに報酬が与えられる慣行があった。よって、再現や複製を行うインセンティブは発生しない<sup>(14)</sup>。複製可能性の危機については、1960年代後半から1970年代前半にかけて問題提起が行われており<sup>(15)</sup>、1970年代後半にはレプリケーション研究が出版されにくいという問題の解決を目的として、“Replications in Social Psychology” というジャーナル<sup>(16)</sup>が発刊された。しかし、3号の発行のみで取り止められている。

近年、ウェブや情報技術の発展により、再現性や複製可能性の検証に必要なデータ共有を容易に実現できるプラットフォームが整備されるようになったこともあり<sup>(17)</sup>、研究の再現性に関する取り組みが進んでいる。出版バイアスを取り除くために、心理学分野を中心として研究計画の事前登録などが実施されている<sup>(18)</sup>。また、科学の報酬システムについては、前述のバッジや学協会による表彰<sup>(19)</sup>等のインセンティブの付与が行われている。

### 4. 研究図書館の役割

再現性・複製可能性について研究図書館が果たす役

割については、セイヤー (Franklin Sayre)<sup>(20)</sup> やステーブス (Vicky Steeves)<sup>(21)</sup> によって述べられている。それらの役割は、大きく教育等の啓蒙活動と情報基盤整備に分けられる。いずれも研究データ管理についての取り組みの中で行われている。

#### 4.1 教育等の啓蒙活動・アウトリーチ

研究図書館は、教材の作成、講義の実施、イベントの開催等によって、再現性・複製可能性についての啓蒙活動を実施している。これらの活動においては、研究者との連携が不可欠である。

具体的には、再現性のある研究を行うための習慣やオープン・スカラーシップについての倫理、再現性のためのリソース等について、ワークショップや授業での講義が実施されている<sup>(22)</sup>。例えば、オランダ・ライデン大学図書館では、“ReprohackNL” というイベントを実施している<sup>(23)</sup>。“ReproHack” は、若手研究者によるイニシアティブであり、参加者は出版済みの研究の再現を試み、著者にフィードバックを提供することにより、再現性に関するツールと課題について学習する。さらに、イベントでの共同作業を通じて、再現性のある研究を目指す研究者の学際的なコミュニティを構築する。“ReprohackNL 2019” は、再現性に関する2件の講演と、論文の再現とその課題の文書化により構成された。このイベントを通して、参加者は再現性のある論文を執筆するためのよい習慣と悪い習慣の両方について学ぶことができ、GitHub や Zenodo といったツールについて経験を積むことができたことが報告されている。

研究者へのアウトリーチに際しては、リエゾン・ライブラリアンを通じて行う各分野の研究者の研究方法や生産するデータについての調査が役に立ったことが指摘されている<sup>(24)</sup>。

#### 4.2 情報基盤整備

研究図書館は、再現性に関する様々な情報基盤整備に携わっている。代表的なものとして、研究データの共有を可能とするリポジトリが挙げられるだろう。しかし、リポジトリで研究データを共有するだけでは、論文の再現性を保証することは難しい<sup>(25)</sup>。

研究の再現の課題として、ファイルの実行環境が挙げられる。同一のファイルを使用していたとしても、異なる実行環境では、異なる結果を出力する、または動作しない可能性がある。この課題を解決するために、米・ニューヨーク大学はReproZip<sup>(26)</sup>というソフトウェアを開発している。ReproZipは、必要なデータファイル、ライブラリ、環境変数、オプションとともに、研究のバンドルを作成する。共有されたバンドルは、

ReproZip を介して、実行することができる。

さらに課題として、データやコードが依存するソフトウェアが廃れてしまうということが挙げられる。この課題を解決することを目的の一つとして、米・イェール大学図書館では、エミュレーションの開発に携わっている<sup>(27)</sup>。

日本国内では、国立情報学研究所が研究データ管理基盤 GakuNinRDM (E2409 参照) 上で研究データを解析するプログラムとその実行環境をパッケージ化し、共有・再利用を可能とするサービスの開発を行っている。サービスにより、研究を容易に再現することが可能となり、さらにその研究に基づいた新たな研究が発展することが期待される<sup>(28)</sup>。

## 5. おわりに

本稿では、再現性・複製可能性を整理した上で、それらについて研究図書館が果たしている役割を述べた。欧米の研究図書館は、研究データ管理の取り組みの中で、啓蒙活動や情報基盤整備を実施している。再現性・複製可能性の危機の根幹には科学の報酬システムがあることが指摘されていた。4章で指摘した役割の他、再現性・複製可能性についてのバッジを、文献のメタデータの一部として流通させ、評価を可能とするような学術情報流通の構築に貢献することも、研究図書館コミュニティには期待されるだろう。NISO の推奨指針では、バッジの実装プロセスを具現化するための新しい委員会の設置が提案されている。研究図書館コミュニティは、これらの動向と協調して取り組みを進めていく必要がある。

- (1) Hensel, Przemysław G. Reproducibility and replicability crisis: How management compares to psychology and economics—A systematic review of literature. *European Management Journal*. 2021, 39(5), p. 577-594. <https://doi.org/10.1016/j.emj.2021.01.002>, (accessed 2022-07-02).
- (2) “UNESCO Recommendation on Open Science”. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en>, (accessed 2022-07-02).
- (3) Sayre, Franklin; Riegelman, Amy. The reproducibility crisis and academic libraries. *College & Research Libraries*. 2018, 79(1), p. 2. <https://doi.org/10.5860/crl.79.1.2>, (accessed 2022-07-02).
- (4) National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. The National Academies Press, 2019, 256p. <https://doi.org/10.17226/25303>, (accessed 2022-07-02).
- (5) Goodman, Steven N. et al. What Does Research Reproducibility Mean?. *Science Translational Medicine*. 2016, 8(341), p. 341ps12. <https://doi.org/10.1126/scitranslmed.aaf5027>, (accessed 2022-07-02).
- (6) Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*. 2015, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>, (accessed 2022-07-02).
- (7) “NISO RP-31-2021, Reproducibility Badging and Definitions”. NISO. 2021-01-28. <https://www.niso.org/publications/rp-31-2021-badging>, (accessed 2022-07-02).
- (8) National Academies of Sciences, Engineering, and Medicine. op cit.
- (9) Romero, Felipe. Philosophy of science and the replicability crisis. *Philosophy Compass*. 2019, 14(11), e12633. <https://doi.org/10.1111/phc3.12633>, (accessed 2022-07-02).
- (10) 以下の文献によると、p値ハッキングは「実験者が納得できる統計的結果が出るまで手段を選ばない不適切な行為」とされる。不適切な行為の具体例としては、意図的または無意識にデータを取捨選択してp値を低くなるように操作する、複数の検定方法のうち一番p値が低い結果のみを報告するといったことが挙げられる。  
三中信宏. 統計学の現場は一枚岩ではない. *心理学評論*. 2016, 59(1), p. 123-128. [https://doi.org/10.24602/sjpr.59.1\\_123](https://doi.org/10.24602/sjpr.59.1_123), (参照 2022-08-01).
- (11) カール・ライムント・ポパー. 科学的発見の論理上. 大内義一, 森博訳. 恒星社厚生閣, 1971, p. 106.
- (12) National Academies of Sciences, Engineering, and Medicine. op cit. Romero. op cit.
- (13) Merton, Robert K. Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*. 1957, 22, p. 635-659. <https://doi.org/10.2307/2089193>, (accessed 2022-07-02).
- (14) Romero, Felipe. Novelty versus Replicability: Virtues and Vices in the Reward System of Science. *Philosophy of Science*. 2017, 84(5), p. 1031-1043. <https://doi.org/10.1086/694005>, (accessed 2022-07-02).
- (15) Ahlgren, Andrew. A Modest Proposal for Encouraging Replication. *American Psychologist*. 1969, 24(4), p. 471. <https://doi.org/10.1037/h0037798>, (accessed 2022-07-02).  
Smith, N. C. Replication studies: A neglected aspect of psychological research. *American Psychologist*. 1970, 25, p. 970-975. <https://doi.org/10.1037/h0029774>, (accessed 2022-07-02).
- (16) Campbell, Keith E.; Jackson, Thomas T. The role and need for replication research in social psychology. *Replications in Social Psychology*. 1979, 1(1), p. 3-14.
- (17) 三浦麻子. 心理学におけるオープンサイエンス—「統計革命」のインフラストラクチャー—. *心理学評論*. 2018, 61(1), p. 3-12. [https://doi.org/10.24602/sjpr.61.1\\_3](https://doi.org/10.24602/sjpr.61.1_3), (参照 2022-07-02).
- (18) 齊藤智. 心理学の研究文化とオープンプラクティス. *情報知識学会誌*. 2021, 31(4), p. 446-451. [https://doi.org/10.2964/jsik\\_2021\\_057](https://doi.org/10.2964/jsik_2021_057), (参照 2022-07-02).
- (19) “ACM SIGMOD Most Reproducible Paper Award Winners”. ACM SIGMOD. <http://db-reproducibility.seas.harvard.edu/awards/>, (accessed 2022-07-08).
- (20) Sayre. op cit.
- (21) Steeves, Vicky. Reproducibility Librarianship. *Collaborative Librarianship*. 2017, 9(2), p. 80-89. <https://digitalcommons.du.edu/collaborativelibrarianship/vol9/iss2/4>, (accessed 2022-07-02).
- (22) Ibid.
- (23) Hettne, Kristina et al. ReprohackNL 2019: How libraries can promote research reproducibility through community engagement. *IASSIST Quarterly*. 2020, 44(1-2), p. 1-10. <https://doi.org/10.29173/iq977>, (accessed 2022-07-02).
- (24) Steeves. op cit.
- (25) National Academies of Sciences, Engineering, and Medicine. op cit.
- (26) Steeves. op cit.
- (27) “Yale Library Emulation Viewer: Home”. Yale University Library. 2022-07-05. <https://guides.library.yale.edu/emulation-view>, (accessed 2022-07-08).
- (28) 藤原一毅. データ解析機能 (GakuNin Federated Computing Services). 2022, 11p. [https://www.nii.ac.jp/openforum/upload/20220601\\_nii\\_rdc4\\_5.pdf](https://www.nii.ac.jp/openforum/upload/20220601_nii_rdc4_5.pdf), (参照 2022-07-29).

[受理:2022-08-16]

Nishioka Chifumi

Reproducibility and Replicability at Research Libraries