

日本の研究者によるプレプリントの活用状況と認識

Utilization and perceptions of preprints by researchers in Japan

池内有為^{*†}, 林和弘[†]

Ui IKEUCHI^{*†}, Kazuhiro HAYASHI[†]

^{*}文教大学, [†]科学技術・学術政策研究所

^{*}Bunkyo University, [†]National Institute of Science and Technology Policy (NISTEP)

あらまし：プレプリントの活用状況と認識を明らかにするために質問紙調査を行った。1,448名の回答を分析した結果、入手は52.1%、公開は20.4%が経験しており、若年層ほど比率が高かった。数学、計算機科学、物理学・天文学分野は入手も公開も盛んであり、化学や医学分野においてもプレプリントが浸透する可能性が示唆された。一方、プレプリントを重視しない研究者の存在も確認された。

キーワード：プレプリント，オープンサイエンス，学術情報流通，学術出版，質問紙調査

1. はじめに

近年、プレプリント（学術雑誌に投稿する予定の査読・出版前の論文草稿）を公開し、迅速に共有する動きが広がっている¹⁾。2020年以降は、COVID-19 / SARS-CoV-2に関する研究の一部がトリガーとなってプレプリントの活用が急増し²⁾、プレプリントの活用を後押しする商業出版社も登場している。一方、査読を経していないプレプリントの信頼性や質の保証などの問題も顕在化している³⁾⁴⁾。こうした状況をふまえて、本研究は日本の研究者によるプレプリントの入手や公開状況、およびプレプリントに対する認識を明らかにすることを目的とする。

2. 方法

2.1. 調査方法と対象

調査方法は、オンラインアンケートシステムを用いた質問紙調査である。調査対象は、科学技術専門家ネットワーク（以下、専門家NW）とした。専門家NWとは、産学官の研究者、技術者、マネージャなど2,000人規模の専門家集団である。メンバーは紹介によって選出され、多分野かつ幅広い年齢層の回答者が含まれることから調査対象とした。

2.2. 調査項目

調査項目は、(1)プレプリントの入手経験（4問）、(2)プレプリントの公開経験（4問）、(3)プレプリントを公開したことがない理由や今後の公開意思（2問）、(4)研究分野と展望（2問）、(5)自由回答（1問）の全13問とした。プレテストは2020年7月28日～7月31日、および8月7日～8月13日に実施して、質問の順序やワーディング等の修正を行った。

2.3. 調査期間と依頼方法

本調査の期間は、2020年8月17日から8月31日とした。調査への協力依頼は、8月17日に専門家NW各位を対象としてE-mailで送信した。リマインダは、未回答者を対象として8月25日に送信した。8月31日以降も回答入力があったため、最終的に9月6日までの回答を結果に含めた。多重回答を防ぐため、回答者ごとに個別URLを作成した。また、回答完了後には再度回答が行えないよう設定した。

2.4. 回答率と回答者の属性

調査は1,914名に依頼し、1,448名から有効回答を得た。回答率は75.7%であった。

回答者の所属は、大学（68.6%）が最も多く、次いで公的機関・団体（17.6%）、企業（13.3%）、その他（0.6%）の順であった。年齢層は40代が最も多く（49.4%）、次いで30代以下（32.4%）、50代（13.3%）、60代以上（4.9%）の順であった。不明の1名（0.1%）は年齢層別の分析からは除外した。研究分野は米国科学審議会（National Science Board）の科学工学指標（Science and Engineering Indicators）の分類から「その他生命科学」を削除して「人文学」を追加した13分野とした。工学（29.8%）が最も多く、次いで生物科学（19.1%）、化学（16.0%）の順であった。回答者数が10名以下であった天文学（10名）は物理学（101名）とあわせて「物理学・天文学」（7.7%）として、人文学（4名）は社会科学（25名）とあわせて「人文学・社会科学」（2.0%）として合計11分野で集計や分析を行った。この偏りは専門家NWの構成によるものであり、特に回答率が低い分野はみられなかった。

3. 結果

3.1. プレプリントの入手経験

プレプリントの入手経験を尋ねた結果、入手経験をもつ回答者は52.1%、なしは46.3%、わからないとする回答者は1.5%であった。所属別に集計すると、公的機関・団体（55.3%）と大学（54.1%）がほぼ同程度であり、企業（37.5%）は比較的少なかった。年齢層別の結果は、30代以下（59.1%）、40代（52.3%）、50代（41.7%）、60代以上（32.4%）の順であった。

図1に分野別の集計結果を示す。数学はすべての回答者がプレプリントの入手経験を有していた。次いで計算機科学（88.1%）、物理学・天文学（86.5%）の順に入手経験をもつ比率が高かった。入手経験をもつ比率が低かったのは、人文学・社会科学（31.0%）、農学（37.4%）、地球科学（37.5%）の順であった。

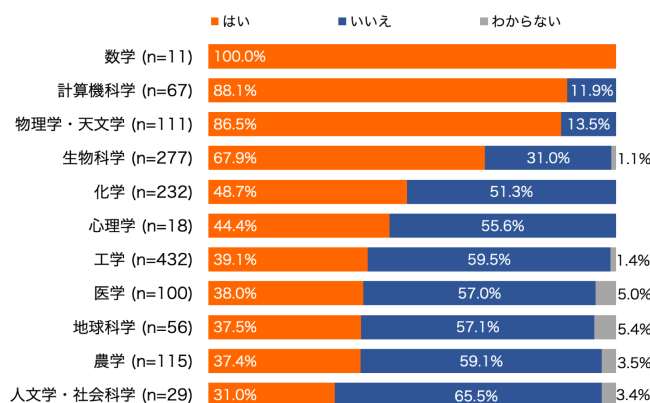


図1. 分野別プレプリントの入手経験

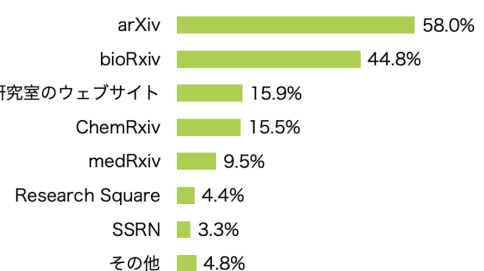


図2. プレプリントの入手先 (n=755)

プレプリントの入手先を明らかにするために、入手経験をもつ回答者755名を対象として複数選択方式で尋ねた。その結果、最も多かったのはarXiv（58.0%）、次いでbioRxiv（44.8%）であった（図2）。

3.2. プレプリントの公開状況

プレプリントの入手経験をもつ回答者755名を対象として、公開経験を尋ねた結果、公開経験をもつ回答者は39.1%、なしは60.7%、わからないのは0.3%であった。全回答者1,448名に対する比率で見ると、公開経験をもつ回答者は20.4%（295名）であった。

所属別に集計すると、公的機関・団体（23.9%）と大学（22.2%）がほぼ同程度であり、企業（6.8%）は比較的少なかった。年齢層別では、30代以下（24.1%）、40代（20.7%）、50代（13.0%）、60代以上（12.7%）の順であった。つまり、入手経験と同様の傾向がみられた。

図3に分野別の集計結果を示す。公開経験をもつ比率が高い順に数学（90.9%）、物理学・天文学（67.6%）、計算機科学（43.3%）であった。公開経験をもつ比率が低かったのは、人文学・社会科学（6.9%）、医学（8.0%）、化学（10.3%）の順であった。入手経験については、化学（48.7%、11分野中5位）や医学（38.0%、同8位）は比較的高かったものの、公開経験は化学（10.3%、同9位）、医学（8.0%、同10位）ともに低かった。

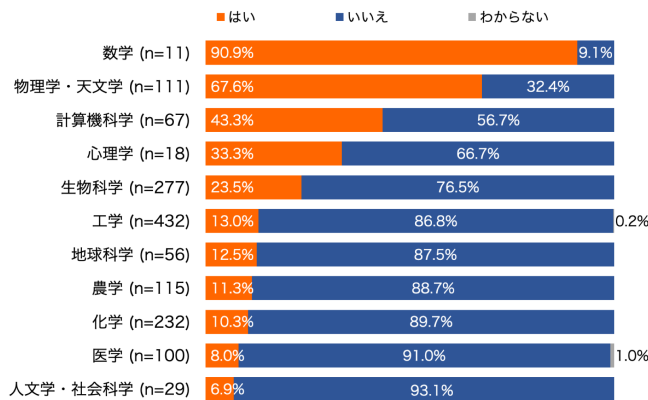


図3. 分野別プレプリントの公開経験

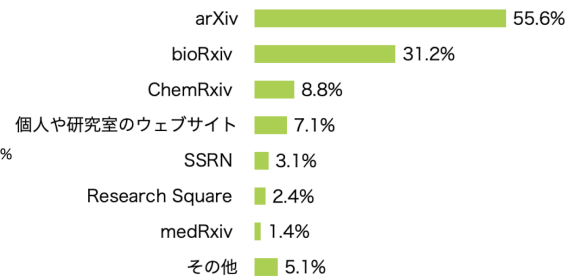


図4. プレプリントの公開先 (n=295)

プレプリントの公開先を明らかにするために、公開経験をもつ回答者295名を対象として複数選択方式で尋ねた。その結果、最も多かったのはarXiv（55.6%）、次いでbioRxiv（31.2%）であった（図4）。medRxivは、入手先として比較的に利用されていたものの（9.5%、7項目中5位）、公開先としてはほとんど利用されていなかった（1.4%、最下位）。

3.3. プレプリントの公開理由と非公開理由

プレプリントの公開理由を、公開経験をもつ回答者295名（全体の20.4%）を対象として複数選択方式で尋ねた。その結果、最も多かったのは「研究成果を広く認知してもらいたいから」（69.0%）、次いで「研究の先取権を確保するため」（65.3%）、「速報性が高いから」（63.6%）の順であった（表1、無回答1名）。

表1 プレプリントを公開した理由 (n=294, 複数回答)

理由	人数	比率
研究成果を広く認知してもらいたいから	203	69.0%
研究の先取権を確保するため	192	65.3%
速報性が高いから	187	63.6%
科学研究や成果実装を推進したいから	53	18.0%
分野・コミュニティの慣習だから	52	17.7%
他の研究者からのリクエストに応じて	49	16.7%
共同研究の契機になる場合があるから	26	8.8%
学術雑誌から投稿の招待を受ける場合があるから	21	7.1%
採用や昇進のため	19	6.5%
採用・昇進以外で業績になる場合があるから	17	5.8%
その他	23	7.8%

プレプリントの公開を敬遠する理由を、今後もプレプリントの公開する意思がないとした回答者 516 名（全体の 35.6%）を対象として複数選択方式で尋ねた。その結果、最も多かったのは「最初に査読誌に投稿したいから」（71.5%）、次いで「プレプリントを公開する必要性を感じないから」（55.2%）、「業績にならないから」（30.6%）の順であった（表 2、無回答 0 名）。

表 2 プレプリントを公開したいと思わない理由（n=516, 複数回答）

理由	人数	比率
最初に査読誌に投稿したいから	369	71.5%
プレプリントを公開する必要性を感じないから	285	55.2%
業績にならないから	158	30.6%
研究分野に適したプレプリントサーバがないから	70	13.6%
その他	64	12.4%

4. 考察

プレプリントは入手、公開ともに若年層の研究者ほど経験をもつ比率が高いことがわかった。所属機関では公的機関・団体や大学の研究者は入手、公開経験をもつ比率が高く、企業の研究者は比較的lowかった。分野による差がみられ、入手や公開先としてよく使われるプレプリントサーバともおおむね一致していた。ただし、入手と公開経験を比較すると、化学と医学は比較的入手経験をもつ回答者の比率が高かった一方で、公開経験はともに低かった（3.2）。プレプリントの公開理由を分野別に集計すると、化学（n=24）、医学（n=8）ともに「採用や昇進のため」を選択した回答者はいなかった。このことから化学や医学分野では、プレプリントを公開することは業績としてあまり評価されていないものの、ChemRxiv や medRxiv に掲載されたプレプリントを利用すること自体は徐々に浸透している可能性が示唆された。

プレプリントを公開する理由は、選択率によっておおむね 3 つのグループに分類できると考えられる（表 1）。すなわち、(1)従来のプレプリントの目的と一致するもの（選択率 60% 台）、(2)近年のオープンな学術コミュニケーションや学術情報流通の流れを汲むもの（選択率 10% 台）、(3)研究業績につながるもの（10%未満）である。(3)に含まれる共同研究の契機や学術雑誌からの投稿の招待、採用や昇進またはそれ以外の業績については、今後、研究者がプレプリントを公開するための新たなインセンティブとなる可能性があると考えられる。

プレプリントを公開したいと思わない理由では、査読誌を優先する、必要性を感じないといった選択肢の選択率が比較的高く、プレプリントを重視しない回答者が一定数存在することが確認された。今後、学術情報流通の慣習や評価の枠組みが変化するならば、プレプリントに対する認識も変化するのではないだろうか。

参考文献

- 1) 林和弘. MedRxiv, ChemRxiv にみるプレプリントファーストへの変化の兆しとオープンサイエンス時代の研究論文. STI Horizon. 2020, vol. 6, no. 1, p. 26-31. <https://doi.org/10.15108/stih.00205>, (accessed 2020-10-20).
- 2) 小柴等, 林和弘, 伊藤裕子. COVID-19 / SARS-CoV-2 関連のプレプリントを用いた研究動向の試行的分析. 文部科学省科学技術・学術政策研究所, 2020, NISTEP DISCUSSION PAPER No. 186, 10p. <https://doi.org/10.15108/dp186>, (accessed 2020-10-20).
- 3) 池内有為. オープンサイエンスの効果と課題—新型コロナウイルスおよび COVID-19 に関する学術界の動向. 情報の科学と技術. 2020, vol. 70, no. 3, p. 140-143. https://doi.org/10.18919/jkg.70.3_140, (accessed 2020-10-20).
- 4) Kwon, Diana. How swamped preprint servers are blocking bad coronavirus research. Nature, No. 581, p. 130-131. <https://doi.org/10.1038/d41586-020-01394-6>, (accessed 2020-10-20).