

## レポート

プレプリントの利活用と認識に関する調査 2020  
— COVID-19 と学術情報流通の現状 —

データ解析政策研究室 客員研究官 池内 有為、室長 林 和弘

## 概要

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）を公開して研究成果を迅速に共有する動きは、COVID-19 を契機として拡大している。そこで科学技術・学術政策研究所（NISTEP）は、今後の学術情報流通政策に資するために、2020年8月から9月にかけて日本の研究者によるプレプリントの利活用の状況と認識に関するオンライン調査を実施した。

調査対象である科学技術専門家ネットワークから得た1,448名（回答率75.7%）の回答を分析した結果、プレプリントの入手経験は52.1%、公開経験は20.4%が有していた。入手、公開のいずれも若年層ほど経験をもつ回答者の比率が高く、所属機関や分野による差もみられた。従来から利用が盛んな数学、物理学・天文学、計算機科学は入手、公開経験ともに高く、化学、医学は公開経験に比して入手経験が高かった。プレプリントの公開理由は「研究成果を広く認知してもらいたい」、「先取権を確保するため」、「速報性が高い」の選択率が高く、わずかながら業績として評価されるという回答もみられた。プレプリントを公開したいと思わない理由は「最初に査読誌に投稿したい」、「必要性を感じない」の選択率が高かった。

キーワード：オープンアクセス（OA）、プレプリント、オープンサイエンス、学術情報流通、オンライン調査

## 1. はじめに

近年、オープンアクセスの潮流のなかでも、プレプリント、すなわち学術雑誌に投稿する予定の査読・出版前の論文草稿をプレプリントサーバなどに公開し、論文出版に先駆けて共有する動きが広がっている<sup>1)</sup>。2020年以降は、COVID-19に関する研究を中心としてプレプリントの活用が急速に増加した<sup>2,3)</sup>。主な背景として、COVID-19は100年に1度というレベルのパンデミックであり国際的な喫緊の課題であること、また、医学分野のみならず多くの分野が課題解決に貢献しうることが挙げられよう。この国際的な課題に対応するために、研究助成機関は国を超えて協働し、学術出版者も呼応した。例えば、「新型コロナウイルスに関連する研究成果とデータを広く迅速に共有する声明」<sup>4)</sup>では、“論文の提出前のデータや前刷りの共有は、学術誌での発表に先駆けた公表とはみなさない”とした上で、“研究成果はデータの利用可能性を明確にした上で、投稿時又は投稿前にプレプリン

トサーバ等で公開する”ことを促した。しかし、研究成果の迅速な共有が進む一方で、査読を経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<sup>5,6)</sup>。

こうした状況をふまえて、科学技術・学術政策研究所（NISTEP）は、今後の学術情報流通政策に資するために、日本の研究者によるプレプリントの利活用の状況と認識に関するオンライン調査を実施した。調査対象は、科学技術予測センターが運営する科学技術専門家ネットワーク（以下、専門家NW）<sup>7)</sup>とした。専門家NWとは、大学、企業、公的機関・団体に所属する研究者、技術者、マネージャーなど、多分野かつ幅広い年齢層を含む2,000人規模の専門家集団である。

調査の概要は次の通りである。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)今後の公開意思とプレプリントを公開したいと思わない理由（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。入手、公開先のプレプリントサーバの例として、1990年代から利用されている arXiv<sup>8)</sup> (物理学から情報学や経済学など多分野に拡大) や SSRN (社会科学から多分野に拡大)、2010年代に登場した<sup>1)</sup> bioRxiv (生命科学)、ChemRxiv (化学)、medRxiv (医学)、ResearchSquare (学際) を挙げた。重複回答を防ぐため、アンケートシステムは回答者ごとに個別 URL を作成した上で、回答完了後には再度回答が行えないよう設定した。調査期間は 2020 年 8 月 17 日から 8 月 31 日として、専門家 NW の 1,914 名に E-mail で依頼した。締切り後も回答の入力があったため、最終的に 9 月 6 日までの回答を結果に含めた。

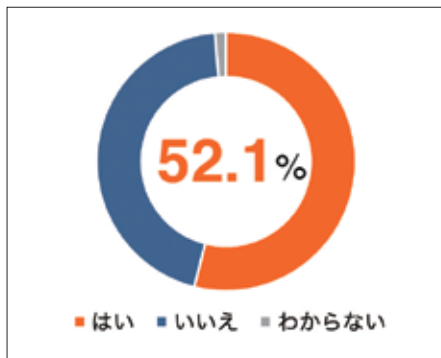
本稿では、1,448 名 (回答率 75.7%) の有効回答を分析した結果から、①プレプリントの入手経験、②プレプリントの公開経験、③プレプリントの公開理由、④プレプリントの未公開理由、⑤プレプリントの展望について紹介する。調査の詳細については、調査資料として公開されている報告書<sup>9)</sup> や発表資料<sup>10)</sup> を御参照いただきたい。

## 2. プレプリントの入手経験

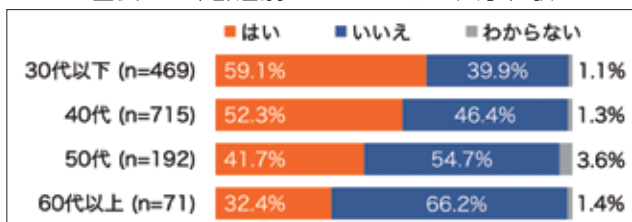
まず、プレプリントサーバ等からプレプリントを入手した経験の有無について尋ねた。結果、入手経験をもつ回答者は 52.1%、もたない回答者は 46.3%、「わからない」という回答者は 1.5% であった (図表 1)。

年齢層別の集計結果をみると、入手経験をもつ回答者の比率が最も高かったのは 30 代以下 (59.1%)、

図表 1 プレプリントの入手経験 (n=1,448)



図表 2 年齢層別プレプリントの入手経験



次いで 40 代 (52.3%)、50 代 (41.7%)、60 代以上 (32.4%) の順であった (図表 2、年齢不明の 1 名を除く)。

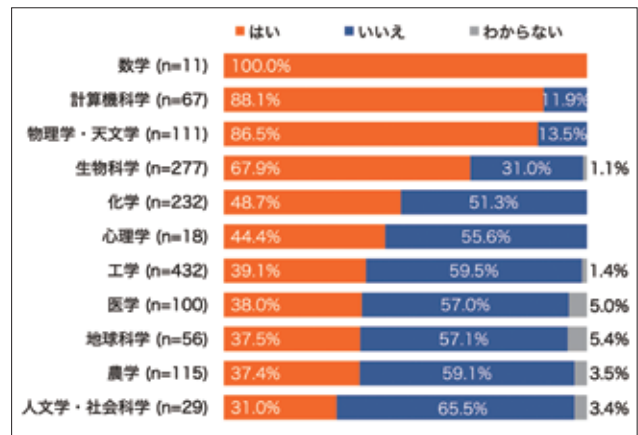
所属機関別の集計結果をみると、入手経験をもつ回答者の比率は公的機関・団体 (55.3%) と大学 (54.1%) が同程度であった。一方、企業 (37.5%) は入手経験をもつ回答者が少なかった。

分野別の集計結果をみると、数学分野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回答者の比率が高かった。最も比率が低かったのは人文学・社会科学 (31.0%) であった (図表 3)。ただし、分野によって回答者数に差があるため、結果を確認する際は御留意いただきたい (以下、同様)。

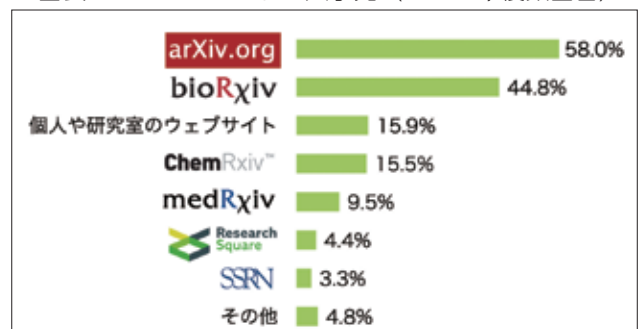
プレプリントの入手経験を有していた回答者 755 名を対象として、プレプリントを入手した際に利用したサーバやサービスを複数選択方式で尋ねた結果、最も多かったのは arXiv (58.0%)、次いで bioRxiv (44.8%)、個人や研究室のウェブサイト (15.9%)、ChemRxiv (15.5%) であった (図表 4)。

プレプリントは査読を経ていないため、速報性が高い一方で内容の信頼性や妥当性の判断は読者に委ねられる。そこで入手したプレプリントの信頼性の判断

図表 3 分野別プレプリントの入手経験



図表 4 プレプリントの入手先 (n=755, 複数回答)



基準を複数選択方式で尋ねた結果、最も多かったのは著者情報（所属機関、職位など）（74.9%）、次いで本文全体（65.9%）、研究手法の確かさ（51.3%）の順であった（図表 5）。「その他」の自由記述では、プレプリントが査読誌や会議録に掲載されたかどうか（7名）、総合的に内容を判断する（5名）、紹介者や他の研究者による評価（4名）、プレプリントの投稿先（3名）といった回答がみられた。なお、プレプリントは参考にする程度で信頼していない（12名）という回答もみられた。

### 3. プレプリントの公開経験

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ところ、公開経験をもつ回答者は 39.1%、もたない回答者は 60.7%、「わからない」という回答者は 0.3% であった（本調査では回答者の負担を軽減するために、プレプリントの入手経験がない回答者は公開経験もないと考えて、入手経験をもつ回答者にのみ公開経験を尋ねた）。プレプリントの入手経験がない回答者を含めた全回答者に対する比率を算出すると、公開経験をもつ回答者は 20.4%、もたない回答者は 79.5%、「わからない」という回答者は 0.1% であった（図表 6）。

年齢層別の集計結果をみると、公開経験をもつ回答者の比率が最も高かったのは 30 代以下（24.1%）、

次いで 40 代（20.7%）、50 代（13.0%）、60 代以上（12.7%）の順であった（図表 7）。つまり、プレプリントの入手経験と同様に、若年層ほど公開経験をもつ回答者の比率が高いという傾向がみられた。

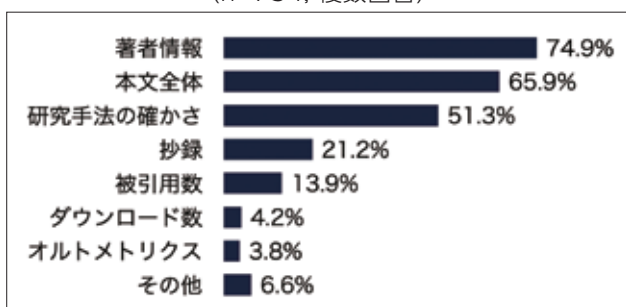
所属機関別の集計結果をみると、公開経験をもつ回答者の比率は公的機関・団体（23.9%）と大学（22.2%）が同程度であった。一方、企業（6.8%）は公開経験をもつ回答者の比率が低く、入手経験と同様の傾向がみられた。

分野別の集計結果をみると、公開経験をもつ回答者の比率が最も高かったのは、数学（90.9%）、次いで物理学・天文学（67.6%）、計算機科学（43.3%）の順であった。最も比率が低かったのは人文学・社会科学（6.9%）、次いで医学（8.0%）、化学（10.3%）の順であった（図表 8）。図表 3 と比較すると、医学や化学分野は、入手経験はやや高いものの（化学 5 位、医学 8 位）、相対的に公開経験は低い傾向がみられた（化学 9 位、医学 10 位）。

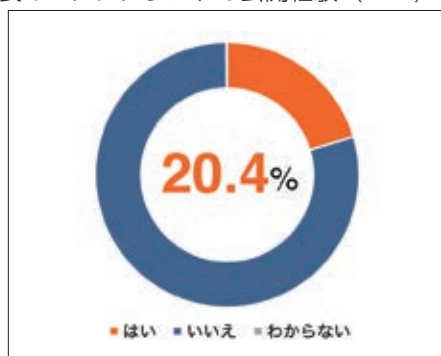
プレプリントの公開先を複数選択方式で尋ねた結果、最も多かったのは arXiv（55.6%）、次いで bioRxiv（31.2%）、ChemRxiv（8.8%）、個人や研究室のウェブサイト（7.1%）であった（図表 9）。プレプリントの入手先とほぼ同様であるが、medRxiv（1.4%）は入手先と比較すると相対的に選択率が低かった。

プレプリントの出版状況を確認するために、プレプリントの公開経験を有していた回答者 295 名を対象として、複数選択方式で出版形式を尋ねた。その結果、

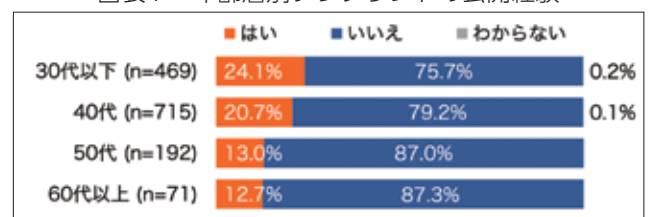
図表 5 プレプリントの信頼性の判断基準 (n=754, 複数回答)



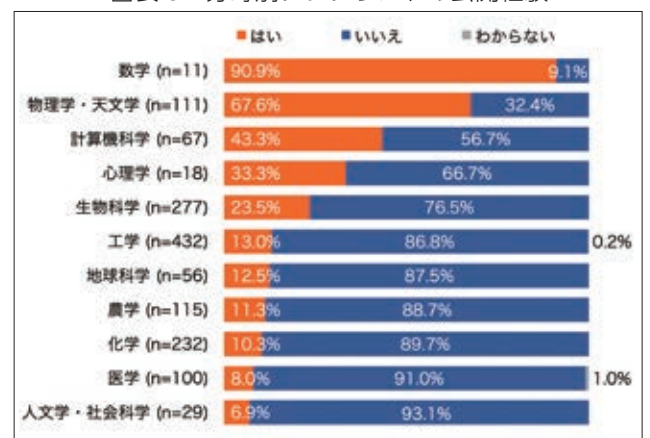
図表 6 プレプリントの公開経験 (n=1,448)



図表 7 年齢層別プレプリントの公開経験



図表 8 分野別プレプリントの公開経験



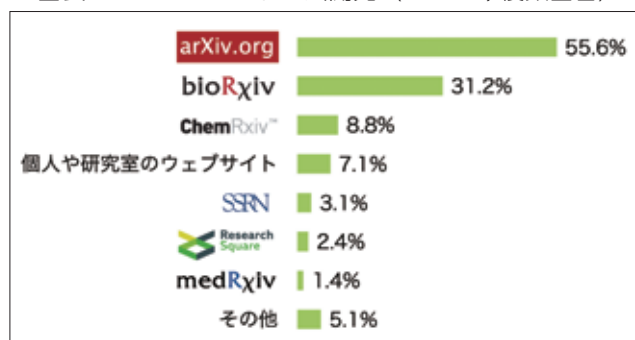


ほとんどの回答者が「論文」(91.5%)を選択していた。わずかながら、「会議録 (conference paper)」(14.2%)や「書籍」(1.4%)という回答もみられた。また、一定数は「プレプリントのみ公開して、出版はしていない」(10.8%)場合があることもわかった(図表 10)。

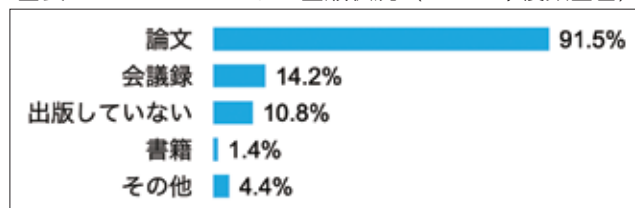
#### 4. プレプリントの公開理由

プレプリントの公開理由を複数選択方式で尋ねた結果、最も選択率が高かったのは「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(図表 11、無回答の1名を除く)。「その他」の自由記述では、雑誌の方針や

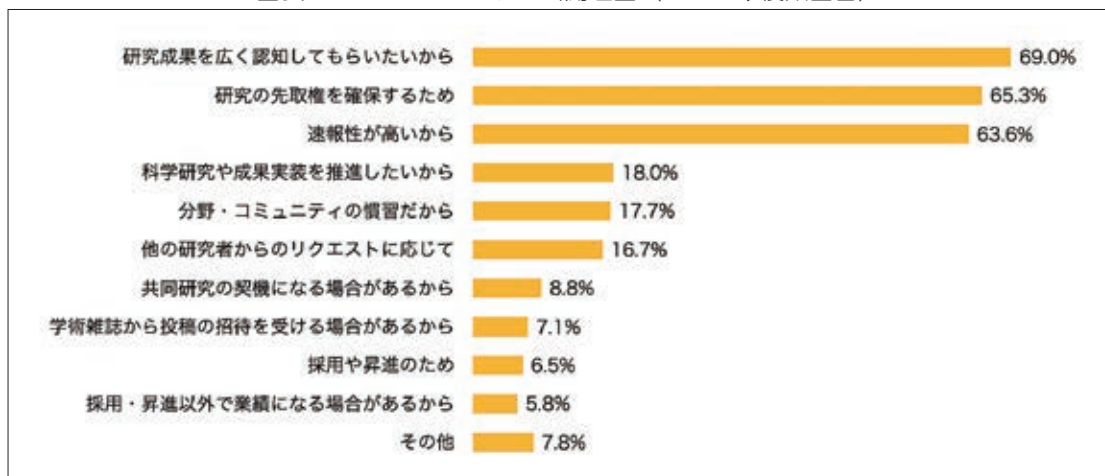
図表 9 プレプリントの公開先 (n=295, 複数回答)



図表 10 プレプリントの出版状況 (n=295, 複数回答)



図表 11 プレプリントの公開理由 (n=294, 複数回答)



仕組み (6名)、引用が可能になるから (5名)、大学から依頼されたから (2名)などがみられた。

#### 5. プレプリントを公開したいと思わない理由

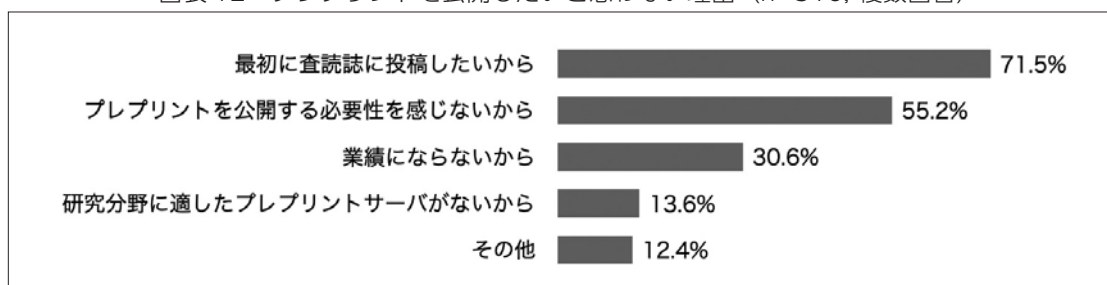
今後プレプリントを公開する意思があるかどうかを、プレプリントの利用経験をもたない、わからないとした回答者、及びプレプリントの公開経験をもたない、わからないとした回答者 1,153名に尋ねた結果、公開意思をもつ回答者は 21.8%、もたない回答者は 48.0%、「わからない」という回答者は 30.3%であった。

プレプリントの公開意思がないとしていた回答者 516名を対象として、プレプリントを公開したいと思わない理由を複数選択方式で尋ねた結果、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(図表 12)。「その他」の自由記述では、査読がなく信頼性が低いこと (24名)や盗用の可能性 (18名)を危惧する意見が多くみられた。また、雑誌によってはプレプリントを公開している場合は投稿を受け付けないとの指摘もみられた (4名)。

#### 6. プレプリントの展望

回答者の分野では、今後プレプリントの利用が進むと思われるかどうかを尋ねた結果、「進むと思う」(32.7%)と「やや進むと思う」(29.3%)の合計は 62.0%であり、6割以上の回答者は利用が進むと考えていることがわかった(図表 13、無回答の8名を除く)。その他の 13名のうち、8名の自由記述は“既に進んでいる”という趣旨の回答であった。

図表 12 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)



図表 13 プレプリント利用の展望 (n=1,440)



## 7. おわりに

本調査では、2020年時点でのプレプリントの入手や公開状況、及びプレプリントに対する認識を明らかにした。1,448名の回答を分析した結果、プレプリントの入手経験は52.1%が、公開経験は20.4%が有していることがわかった。入手、公開のいずれも若年層ほど比率が高く、所属機関別では公的機関・団体と大学が同程度であり、企業はやや低かった。

分野別の状況を確認すると、1990年代からプレプリントサーバが登場し、活用が盛んであるとされている数学、物理学・天文学、計算機科学は、入手、公開経験をもつ回答者の比率が高かった。一方、人文学・社会科学、農学、地球科学は入手、公開経験をもつ回答者の比率が相対的に低かった。ChemRxivやmedRxivのプレプリント登録数が増えている中で、化学や医学は入手経験をもつ回答者の比率は高かったものの、公開経験をもつ回答者の比率は相対的に低かった。プレプリントサーバの登場やCOVID-19に関する研究成果の迅速な共有を契機として、まずは利用が広がっている可能性が示唆された。

論文は掲載雑誌及び査読によって一定の信頼性が担保されている一方で、プレプリントについては利用者が信頼性を判断する必要がある。プレプリントを

入手した際の信頼性の判断基準としては、「著者情報(所属機関、職位など)」、「本文(全文)」、「研究手法の確かさ」などが選ばれていた。また、自由回答では、査読がないため参考にする程度で信頼していないという趣旨の回答も多くみられた。

プレプリントの公開理由は、「研究成果を広く認知してもらいたいから」、「研究の先取権を確保するため」、「速報性が高いから」の比率が高かった。また、わずかながらプレプリントも業績として評価されている場合があることもわかった。プレプリントを公開しようと思わない理由は「最初に査読誌に投稿したいから」、「公開する必要性を感じないから」、「業績にならないから」の選択率が高かった。

今後も継続的に調査を行い、プレプリントの利活用や受容が一般化するのか、あるいは、分野等によってどのように変化するのかを明らかにしていきたい。また、調査対象者を拡大することによって、より正確に日本の研究者の状況を捉えることを目指したい。

## 謝辞

調査及びプレテストに御協力を賜った皆様に心よりお礼申し上げます。

- 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 2021-04-18).
- 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 2021-04-18).
- 3) Else, Holly. How a torrent of COVID science changed research publishing — in seven charts. Nature, 2020, no. 588, p. 553. <https://doi.org/10.1038/d41586-020-03564-y>, (accessed 2021-04-18).
- 4) “日本医療研究開発機構 (AMED) は新型コロナウイルスの流行に対処するため、新型コロナウイルスに関連する研究成果とデータを広く迅速に共有する声明 (令和 2 年 1 月 31 日) に署名しました”. 国立研究開発法人日本医療研究開発機構. 2020-02-03. <https://www.amed.go.jp/news/topics/20200203.html>
- 5) 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](https://doi.org/10.18919/jkg.70.3_140), (accessed 2021-04-18).
- 6) 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 2021-04-18).
- 7) “科学技術専門家ネットワーク”. 文部科学省科学技術・学術政策研究所. <http://www.nistep.go.jp/activities/st-experts-network>, (accessed 2021-04-18).
- 8) 三根慎二. 学術情報メディアとしての arXiv の位置づけ. Library and information Science. 2009, no. 61, p. 25-58. [http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara\\_id=AN00003152-00000061-0025](http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00003152-00000061-0025)
- 9) 池内有為, 林和弘. プレプリントの利活用と認識に関する調査. 文部科学省科学技術・学術政策研究所, 2021, NISTEP RESEARCH MATERIAL No.301, 94p. <https://doi.org/10.15108/rm301>
- 10) 池内有為, 林和弘. “日本の研究者によるプレプリントの活用状況と認識”. 情報メディア学会第 22 回研究大会. オンライン開催, 2020 年 11 月 7 日.