

# オープンサイエンス時代の言語系研究と教育 —プレプリントの公開をめぐる—

横山詔一（国立国語研究所研究系）

石川慎一郎（神戸大学大学教育推進機構／国際文化科学研究科／数理・データサイエンスセンター）

責任著者：横山詔一（yokoyama@ninjal.ac.jp）

## 要旨

オープンサイエンスの推進に必要な不可欠なプレプリントの公開に着眼し、それが言語系研究の成果発表や大学院教育の在り方にどのような影響を与えつつあるのか、また、そこにどのような可能性と問題があるのかを概観する。第2節では日本語対応プレプリントサーバーである Jxiv が誕生した背景について述べる。第3節ではプレプリントの倫理面における諸問題を取り上げる。著作権法等に関する法律的側面には立ち入らず、研究者同士の信頼関係に影響するかもしれない心理的側面について議論する。第4節ではプレプリントサーバーを言語系大学院教育で活用する意義を示し、授業実践の例を紹介する。そして、第5節では研究者 SNS である ResearchGate を公刊済み論文のプレプリントの公開プラットフォームとして使用する可能性と課題について概観し、機関レポジトリとの関係についても言及する。最後に、第6節において本稿の議論を総括する。

## キーワード

言語系研究、オープンサイエンス、プレプリントサーバー、Jxiv、クリエイティブ・コモンズ・ライセンス、大学院教育、研究者 SNS、ResearchGate、機関レポジトリ

## 1. はじめに

オープンサイエンスとは、[日本学術会議 \(2020\)](#) によると「ICT（情報通信技術）を活用して科学を変容させること」とされている。より詳しい別の説明には「研究者のような専門家だけでなく非専門家であっても、あらゆる人々が学術的研究や調査の成果やその他の発信される情報にアクセスしたり、研究活動に多様な方法で参加したりできるようにするさまざまな運動のこと」と述べられている（[Wikipedia](#)、2022年8月8日アクセス）。オープンサイエンスの出発点は、国際ヒトゲノム計画で読み取ったヒトゲノム配列は、24時間以内にすべて公開するというバミューダ原則（1996）にある（有田、2021）。

本稿は、こうしたオープンサイエンスの推進に必要な不可欠なプレプリントの公開に着眼し、それが言語系研究の成果発表や大学院教育の在り方にどのような影響を与えつつあるのか、また、そこにどのような可能性と問題があるのかを概観する。学術出版におけるプレプリントの定義には曖昧さも残るが、本稿では、米国情報標準化機構（NISO）と学協会出版社協会（ALPSP）の2008年の勧告を解説している[尾城 \(2021\)](#)などを参考にして、投稿前ないし投稿時点における著者オリジナル版（査読による修正が一切入っていないもの）と、査読過程版（一次査読などの意見で部分的な修正が入っているもの）をプレプリントとする。なお、受理（accepted）が決定した著者最終版（出版社の編集などが入っていないもの）はポストプリントと呼ばれ、プレプリントとは区別される。

プレプリント公開は、ふつう、ジャーナルなどでの論文の掲載に先立って行うが、ジャーナル掲載後に行う場合もある。前者は、投稿前に論文を公開して読者からの助言で投稿論文の質を上げる、価値ある知見を公益のためにいち早く世に示す、知見の先取権を主張する、といった目的でなされる。後者は、ジャーナル出版社の著作権を侵害しない形で自身の論文へのアクセス手段を複数化し、論文の購読数や引用数をあげる目的でなされることが多い。

こうした目的により、プレプリントの公開先も変わってくる。掲載前論文は、arXiv や Jxiv といった公共性を持つプレプリントサーバーに公開されるのが一般的である。また、掲載後論文は、ResearchGate や Academia.edu などの研究者 SNS や、機関レポジトリなどに公開することができる。このほか、ジャーナルなどへの投稿を前提にしない短報や研究メモを公開したい場合は、研究者 SNS に加え、科学技術振興機構が運営するデータベース型研究者総覧である Researchmap の「研究ブログ」や「資料公開」ページなども活用できる。

研究者は、運営元の信頼性（公共か民間か、営利か非営利か）、予想される購読者属性（研究者か一般読者か）、予想される購読者数、公開に伴う制約の度合い（公開範囲が自由に選択できるか、公開済み論文を撤回・差し替えできるか）などを総合的に判断し、自身の目的にとって最適のプラットフォームを選んで公開を行う必要がある。

ただ、これまで、こうしたサービスを積極的に活用してきたのは自然科学系の研究者が多く、人文系研究者、とくに、言語学・応用言語学・言語教育学（本稿では以下、これらを言語系研究と総称する）分野の研究者がこうしたサービスを利用した経験は相対的に少ない。そこで本稿は、二名の著者の最近の体験をふまえ、「投稿前論文のプレプリントの Jxiv での公開」と「公刊済み論文のプレプリントの ResearchGate での公開」に焦点を絞り、各々の概要や意義、問題点、また、教育的活用について論じる。ただし、著作権関連の法的な議論や、各種の論文公開プラットフォームの総合的な概観は、著者の守備範囲を超えるため、本稿では扱わない。本稿は、あくまでも、著者個人の体験の範囲に基づく報告であることを最初に断っておきたい。

以下、第2節では Jxiv が誕生した背景について述べる。第3節ではプレプリントの倫理面における諸問題を取り上げる。ただし、著作権等に関する法律的側面には立ち入らず、研究者同士の信頼関係に影響するかもしれない心理的側面について議論する。第4節ではプレプリントサーバーを言語系大学院教育で活用する意義を示し、授業実践の例を紹介する。第5節では ResearchGate を公刊済み論文のプレプリントの公開目的で使用する可能性と課題について概観し、あわせて、大学図書館が運営する機関レポジトリとの関係についても考察する。最後に、第6節において本稿の議論を総括する。

## 2. 日本語対応プレプリントサーバー誕生の背景

### 2. 1 Jxiv の登場

日本でも医学研究者はプレプリントサーバーをよく利用している。医学研究者の多くは、論文を学術誌に投稿することを前提として、プレプリントサーバーに先に、あるいは同時に投稿し、最新の情報を研究者コミュニティに提供・共有しつつ、査読プロセスを進めるようにしている（柴田、2022）。新型コロナウイルス感染症に対処する研究が人類全体の喫緊の課題となったこともあって、プレプリントサーバーの活用は学術研究の世界で急拡大を続けている。最新の研究成果をプレプリントサーバーによって世界中の多くの研究者が迅速に共有することで、研究スピードを一気に加速できるからである。その結果、プレプリントが医学研究には必要不可欠の存在である

ことがますます強く認識されるようになった。

ひるがえって、言語系研究者とプレプリントサーバーの関係に眼を向けてみよう。これまで、日本の言語系研究者のうちプレプリントサーバーに投稿した経験を有する人はきわめて少ないのではなかろうか。その主要な理由のうち2つを示す。1つ目は、プレプリントという概念を言語系研究者の大部分が（無関心で）知らないからである。2つ目は、日本語で投稿できる（本格的な）プレプリントサーバーが存在しなかったからである。

この状況を打破するために、JSTは日本語でも投稿できるプレプリントサーバーJxiv（ジェイカイブ）を2022年3月24日から稼働させた。JSTは、国立研究開発法人 科学技術振興機構（Japan Science and Technology Agency）の略称で、研究費にかかる事業にも参画している組織である。ちなみに、Jxivをジェイカイブと読むのは、Xがギリシャ語の「カイ」であることに由来するようである。

## 2. 2 Jxivに関するNatureのニュース記事について

Jxiv稼働から3ヶ月も経過していない2022年6月6日に学術誌NatureがJxivを紹介する [Singh Chawla \(2022\)](#) の記事を掲載した。記事は、Jxivの可能性や特長として、日本語論文を投稿できる、プレプリントの共有が促進される、日本の研究を世界の学术界に発信できる、既存のプレプリントサーバーが扱ってこなかった歴史・ビジネス・経営・言語・学際科学分野の日本における研究に光を当てることができる、といった点を指摘している。その一方で、日本人研究者は公の場での他者の研究への批判を控える傾向が強いため、一般の査読論文に比べ、プレプリントで公開された論文の質が保証されない懸念もあると述べている。

また、このNatureニュース記事の最後は、熊本大学の発生学者であるGuojun Sheng氏による「Jxivは普及すると考えている。特に、資金提供機関が、将来的に資金を提供する研究者にJxivの使用を義務付けるようになれば、なおさらである」という見解で締めくくられている。

Jxivの登場が学術研究の環境にどのように影響するのかを現時点で見通すことは難しい。しかし、大学の評価基準に「オープンサイエンス推進に取り組む姿勢・熱意の指標」の一つとして「所属教員がプレプリントで公開した論文の本数」といった尺度が近いうちに新たに付加される確率は低くはないようにも思える。当然のことながら、プレプリントの本数をもって「オープンサイエンス推進に取り組む姿勢・熱意の指標」とすることについて、まったく問題がない訳ではない。しかし、これ以外に、より有効で簡便な指標はなかなか見つけられそうにない。この指標は、この種の評価としては「公正、客観的で妥当性と信頼性が高い」ものだと判断しても大過ないと言えそうである。

## 3. プレプリントの倫理面に関する諸問題

### 3. 1 投稿者（著者）はプレプリントを削除できない

Jxivに投稿すると、形式面が整っていれば、プレプリント公開まで原則48時間以内で進む。これは投稿者には都合がよいように見えるが、以下のような点に注意すべきである。

- プレプリントにはDOI（Digital Object Identifier）が必ず付与される
- プレプリントは基本的に削除できない
- プレプリントを削除する権限はJSTが有する

プレプリントが公開されると投稿者の意思で削除することはできない。よって、投稿者は、内

容面は当然のことながら、倫理面にも十分配慮し、トラブルを未然に防止する方策を慎重に検討する必要がある。ところが、プレプリントの倫理面に関していかなる問題が予想されるのか、その回避策として具体的に何があるのかを、言語系研究者の大部分はほとんど知らないのが現実であろう。

以下、倫理面にフォーカスして、何が問題なのかを順番に検討していく。ここでは著作権法等にもとづく法律的側面の議論はしない。あくまでも研究者同士の信頼関係に影響すると予想される心理的側面について考えることを強調しておく。

### 3. 2 プレプリントを投稿できる言語系学会査読誌は希少

日本の主要な言語系学協会が刊行する査読誌の中で、すでにプレプリントが公開されている論文の投稿を二重投稿と見なさいと規定に明示しているのは、第一著者が調べた範囲では一誌のみであった（2022年5月時点）。紀要も同じような状況で、それを明示的に認めているのは『国立国語研究所論集』のほか、ごくわずかの大学（院）紀要のみのものである。よって、現時点では言語系論文をJxivに投稿すると大部分の学会査読誌には投稿できなくなる。また学会査読誌に投稿できたとしても、不採択となることも念頭に置いて対応策を考えておく必要がある。

なお、プレプリントを投稿できる言語系学会査読誌が希少であるという問題は、今後の世界的なオープンサイエンス化の流れのなかで徐々にあるいは急速に解消されていくと予想する。

### 3. 3 どのクリエイティブ・コモンズ・ライセンスを選択するのか

ここで登場する専門用語とその短い説明をまず示す。

- クリエイティブ・コモンズ・ライセンス (creative commons (CC) licenses) : インターネット時代のための新しい著作権ルールで、作品を公開する作者が「この条件を守れば私の作品を自由に使ってよい」という意思表示をするためのツール
- CC BY : 出典情報さえ示せば、第三者による改変や営利目的の使用も許すもので、クリエイティブ・コモンズ・ライセンスのなかで制約が一番ゆるく、著作権法上のパブリックドメインに近い
- ND : No Derivative Works の略語で改変禁止を示す
- NC : No Commercial Use の略語で非営利（営利目的の利用禁止）を示す

#### (1) *Nature* に掲載された論文は CC BY がスタンダード

Jxiv に投稿する際は、論文のクリエイティブ・コモンズ・ライセンスを予め決める必要がある。クリエイティブ・コモンズ・ライセンスは6種類あり、Jxiv 投稿時にそのうちの1つを必ず選択しなくてはならない。また、投稿時に選択したクリエイティブ・コモンズ・ライセンスは永久に変更できない。ここで投稿者はどれを選ぶか大いに迷うことになる。クリエイティブ・コモンズ・ライセンスの6種類について、どのような違いがあり、今後いかなる事態が起こりうるのか皆目見当がつかないからである。

このような問題に直面した場合、オープンサイエンスの世界的動向を探り、それを参照・吟味しながら意思決定をはかることが重要であろう。とりわけ、ヨーロッパの著名な学術出版社が取り組んでいる学術誌のオープンアクセス化が参考になる。たとえば、*Nature* に掲載された論文のクリエイティブ・コモンズ・ライセンスはCC BY がスタンダードになっている。

#### (2) 学会査読誌の CC 表示について

Jxiv に投稿を目指す研究者は、Jxiv で公開されるプレプリントを投稿したい学会査読誌の投稿

規定にクリエイティブ・コモンズ・ライセンスに関する記述があるかを確認する必要がある。そこで念頭におくのは次のような事態を回避することである。仮想例を時系列で示す。

1. 研究者の A 氏が Jxiv に論文を投稿した。投稿時にクリエイティブ・コモンズ・ライセンスの CC BY を選択し、プレプリントが公開された。
2. プレプリント公開と同時に A 氏はプレプリントを認めている学会査読誌に投稿した。
3. プレプリント公開から数日後、第三者が内容を一部改変したうえで営利目的の利用を開始した。
4. プレプリント公開から半年後、査読を通過してプレプリントと瓜二つの内容の論文が公刊されたのだが、学会の判断で CC BY に NC と ND が追加され、CC BY-NC-ND となった。

この場合、さまざまな不都合な事態が生じることが懸念される。まず、学会査読誌で論文が公刊されたのは第三者が改変をおこなった後である。営利目的の利用もすでに開始されている。このようなケースで、学会側は、改変禁止と、営利目的の利用差し止めを主張できるのであろうか。第三者側は「あくまでも CC BY のプレプリントを改変し、営利目的に利用したのであって、学会側の主張は的外れだ」と反論するであろう。要するに、プレプリントが CC BY であれば、学会側が追加した NC と ND は実質的に無効になるのではないか。

Jxiv に投稿を目指す研究者は、これらのことを意識しながら、投稿に先立ってクリエイティブ・コモンズ・ライセンスの件を学会に問い合わせ、投稿しても大丈夫か確認すべきである。

### (3) 共著論文の CC 表示には慎重な意思決定が必要

共著論文で CC BY に ND (改変禁止) などの制約を追加したクリエイティブ・コモンズ・ライセンスを選択する場合は、以下のメリットとデメリットの両面を明確に認識したうえで、共著者全員の合意に基づいて慎重に意思決定を行うのがよいだろう。

- **メリット** : Jxiv 投稿時点では予想できなかった問題が将来生じた場合に備えて、論文の著者が権利を主張する余地を残しておく。ND や NC をつけておかないと、今後、なんらかの事情で改変や商業利用を差し止めたくなった場合に対抗策が打てなくなる。これは、著作権法上の安全策と言えよう。
- **デメリット** : ND をつけると共著者間の信頼関係が壊れる原因になりかねない。以下の仮想例で、今後予想される典型的なトラブルを示す。

A 氏と B 氏が共著で Jxiv に投稿し、CC BY-ND でプレプリントが公開された。その後、A 氏は当該プレプリントの内容を単独で要約し、ある企業が主催した講演で利用した。A 氏はこの程度の短い要約であれば改変にはならないと判断し、事前に B 氏に相談することはしなかった。ある日、B 氏はその要約をたまたま眼にする機会があった。B 氏は A 氏の要約を慎重に検討した結果、当該プレプリントの内容を正確に伝えきれていない (いささか適切とは言えない) 表現があるとの印象を持った。B 氏は心のなかで「これは改変にあたるのではないか、弁護士に相談してみよう」とつぶやいた。

この後、さまざまな展開が想像できるだろうが、未然に回避したい状況であることは衆目の一致するところであろう。上記の仮想例のように、CC BY に ND などの制約を追加すると、第三者に対してだけでなく共著者全員にもその制約を永久に課すことになりかねない。そのことが、著作権法等にもとづく法律的側面の議論とは関係なく、共著者同士の信頼関係にダメージを与える危険性がある。このリスクを未然に回避するシンプルな方法の一つが CC BY を選択することだと考える。

## 4. 大学院教育で Jxiv を活用する意義と授業実践例

### 4. 1 大学院教育で Jxiv を活用する意義

先に紹介した *Nature* の Jxiv に関するニュース記事は無視できない存在感がある。今のうちから言語系の大学院教育の場においても、プレプリントサーバーに投稿する方法や注意点を教授することには一定の意義があると考えてよいだろう。大学院生が研究者のポストに就くためにも、またそのポストよりも条件の良いポストに移るためにも、(医学研究者と同様に) プレプリントで論文を公開することを念頭に置いて計画的に努力することが求められる時代が、目前に迫っていると予測するからである。

そこで、以下では Jxiv を用いた授業を紹介する。東京大学大学院総合文化研究科言語情報科学専攻の 2022 年夏 semester 開講科目の一つに「言語変化のデータサイエンス基礎」があった(本稿の第一著者である横山が担当)。その授業ではデータ解析手法としてロジスティック回帰分析や分散分析などを取り上げ、理論面の解説に加えて具体的な解析手順を説明した。さらに、学術論文においてロジスティック回帰分析の結果をどう記述するかを具体的に示すため、ロジスティック回帰分析を用いたプレプリント ([横山・前田・高田・相澤・野山・福永・朝日・久野, 2022](#)) を教材の一部として利用することにした。その具体的な流れを紹介する。

### 4. 2 授業実践の例

#### (1) 他人(教員)が投稿した論文の査読プロセスを受講生が観察する

先に引用した [柴田 \(2022\)](#) には「医学研究者の多くは、論文を学術誌に投稿することを前提として、プレプリントサーバーに先に、あるいは同時に投稿し、最新の情報を研究者コミュニティに共有しつつ、査読プロセスを進めるようにしている」とある。プレプリントサーバーに投稿するという事は、瓜二つの論文を学術専門誌に投稿して査読を受けることを第三者に開示することになるばかりか、投稿論文の内容もネットで公開される。

あくまでも一般論であるが、学会査読誌に投稿するからには、論文の著者は「不採択のリスク」も覚悟している。著者は、不採択になった場合に備えて、学会査読誌に投稿したという事実はもちろん、その投稿論文の内容も第三者に知られたくないという意識を持つ。つまり、第三者に知られぬよう「ひっそりと学会査読誌に投稿する」という時代が続いてきたのである。しかし、この状況を根底から破壊するイノベーションの波が押し寄せてきた。それがプレプリントサーバーである。学会査読誌に「ひっそりと投稿する」時代はそのうち終焉を迎えるであろう。

このことをふまえて、プレプリント ([横山ほか, 2022](#)) を 2022 年 5 月 20 日にある学会査読誌に投稿した。そして、その事実を「言語変化のデータサイエンス基礎」受講生全員に公開した。プレプリントが公開されたことで、他人(この場合は教員)が学会誌に投稿した論文と同じ内容の PDF ファイルを受講生全員で共有することが実現した。この授業の受講生は、教員の論文が学会の査読プロセスをどのように進んでいくのか、その一部とはいえ、毎週の授業でなされる教員の報告を通じて観察できる環境が整ったことになる。教員の投稿論文が学会査読誌に採択されるのか、はたまた不採択となるのか。受講生の興味や学習動機は自然に高まるであろう。

授業では、投稿論文に示されたデータを用いて受講生がロジスティック回帰分析を行い、論文に記載された解析結果と数値が一致するか検算する課題も出された。これは、受講生が査読者になったつもりで教員の投稿論文を子細にチェックし、自分なりの評価を出すということである。大学院教育でこのような査読者シミュレーションを可能にするツールとしてプレプリントサーバ

一は使いやすい。

## (2) 研究倫理について具体的に考える

教員は受講生に対して、プレプリントサーバーに投稿する際はターゲットとする学会査読誌で二重投稿とみなされないかを投稿者自身が慎重に確認すべきであることを解説し、あわせて受講生自身が常にさまざまな学会査読誌の投稿規定をチェックするように促すのがよいだろう。

次に、クリエイティブ・コモンズ・ライセンスについて、教員による解説が必要である。Jxiv 投稿においては、投稿論文の内容や形式、そして共著者との関係性などを十分に考慮して、どのクリエイティブ・コモンズ・ライセンスを選択するかを決めることになるのだろうが、抽象的な説明に終始しては受講生に説明が十分に伝わらない恐れがある。実例を示した方が教育効果の増進が期待できるだろう。そこで、受講生に[横山ほか \(2022\)](#) のプレプリントを 2022 年 5 月 26 日 (木) の授業でダウンロードさせた (受講生全員が教室に PC を持参しネットに接続)。論文本体のほか、論文本体の後ろに付した「倫理チェック」リストについて教室でディスカッションした。なお、そのディスカッションの内容については、受講生の事前承諾を得ていないため、今回は開示できない。

## 5. オープンサイエンスと研究者 SNS

### 5. 1 プレプリント共有プラットフォームとしての RG

上記で触れた Jxiv に加え、研究者はまた、いわゆる研究者 SNS でも自身の研究を公開することができる。国際的な研究者 SNS としては Academia.edu や ResearchGate (RG) などが知られているが、ここでは第二著者が使用経験のある RG に限定して概観する。

RG は 2008 年に米国で創始され、現在は本拠をドイツに置く。2015～2017 年に海外で実施された複数の調査によると、同種の SNS サービス中、RG の利用者数は最多であったとされる ([坂東、2018](#))。その後も RG の利用者は増えており、本稿執筆時点で、大学等の所属機関が発給した公式メールアドレスによって登録された会員数は 2,000 万人を超え、1 億 3,500 万ページを超える論文が公開されている。後述する問題はあるものの、国境を超えた研究者のネットワーキングや、オープンサイエンスの推進に RG が大きな貢献を果たしてきたことは否定できないであろう。

RG は、同じ分野の会員をフォローしたり、会員間で論文のリクエストを行ったり、特定のテーマについて他の会員と議論したりすることが可能であることから、少なくとも初期の段階では「研究者のための SNS」とみなされていた。しかし、ネットワーキング機能の使用は必須でなく、次第に、プレプリントの共有プラットフォームとしての性質が強まってきたように思える。[坂東 \(2018\)](#) も、海外での調査結果をふまえて、各種の研究者 SNS が「誕生当初は他研究者とのネットワーキングや自身のオンラインプロフィールとして活用されることが主流であったが、近年は論文共有の場として使われている」と述べている。

### 5. 2 RG の利点

ここでは、第二著者の経験の範囲で、言語系研究者が公刊済み論文のプレプリントを RG で公開するメリットとして 5 点を指摘したい。

#### (1) 幅広い読者へのリーチ

これは、国際誌などに掲載された論文についても、学内紀要や各種の報告書などに掲載された論文についても当てはまる。前者は、通例、購読費用が発生するため、後者は認知度が限られる

ため、それぞれ読者は限定される。しかし、RGにプレプリントを公開すれば、長期にわたって圧倒的多数の潜在的読者の目に触れる。第二著者は2014年にRGに登録し、以降、自身の公刊済み論文のプレプリントをRGに公開しているが、それらは当初の予測を上回って広く読まれている。RGのようなプラットフォームがなければ、学内紀要に出した論文などが、海外を含めた幅広い読者に読まれることはおそらくなかっただろうと思われる。

## (2) 論文ごとの受容状況の把握

会員は、自身が公開した論文の各々について、購読数（概要・図・表などを含め、論文の一部が表示された回数）、完全購読数（論文の全体が表示されたり、ダウンロードされたりした回数）、引用数などの統計データにアクセスできる。これにより、自身の研究の中でどのようなものが他の研究者の関心を引いているのか、主としてどういう国・地域の研究者に読まれているのか、過去において購読数・引用数がどう変化してきたのか、といった点について、客観的な把握が可能になる。

一般に、研究者は、査読の段階までは、査読者にどう読まれるかを強く意識する。しかし、いったん論文が出版されると、それが誰にどう読まれているか知る手段はほとんどなく、関心を向けることも少ない。この意味において、RGが提供する論文ごとの詳細な受容データは、研究者が自身の研究を振り返り、必要に応じて研究の方向性を再考するきっかけとなりうる。

## (3) 刊行済み論文の長期的管理

論文へのアクセスを増やすだけなら個人サイトでの公開も考えられる。しかし、異動や退職でサイトが閉鎖されることもあり、論文公開用のURLを長く保持すること容易ではない。RGでは、一度公開されたものは、（運営会社がサービスを継続する限り）永続的にアクセスでき、URLの同一性も保たれる。また、RGでDOIを取得し、論文をより体系的に管理することもできる。

RGには、Jxivのように、一度公開した論文を取り消せないという規定はない。このため、後で誤植が見つかった場合など、新しい版に差し替えることもできる。RGでDOIを取得した論文の差し替えはできないが、当該論文を削除し、修正版を新たに公開してDOIを取り直すことが可能である（この件については [t.ly/IF1](#) を参照されたい。以下、本節内での参照URLはすべて短縮表示で示す）。

## (4) 研究のモチベーションの維持

著書の刊行や定評ある国際誌での採択は、言うまでもなく、大きな研究上のモチベーションとなりうるが、言語系研究分野において、そうした業績を毎年出し続けていくことは容易ではない。この意味において、そうした業績以外であっても、自身の研究が誰かに読まれているという事実を知ることは、研究者が研究を続けていく上で大きな励みとなる。RGでは、定期的な購読数の通知に加え、新規の引用があるたびにメールが届き、論文が読まれていることを実感しやすい。自然科学系に比べ、研究に注目が集まりにくい言語系の研究者にとってもこの点は有益である。

一例として、第二著者は2011年公刊の論文で英単語の綴りに関するきわめて周縁的な問題を扱った。この論文は公刊後もほとんど反応がなかったが、2015年にRGでプレプリントを公開してから少しずつ読まれるようになり、刊行後10年を経た2021年になって、フィンランドの研究者によって初めて引用された。こうした経験は、研究者が地味な研究を粘り強く続けていく支えとなってくれる。

この意味において、RGのようなプラットフォームの意義は、自然科学系と人文系でいくぶん異なるのかもしれない。周知のように、自然科学系では、トップジャーナルで公開されるような

「査読論文しか評価しない」という傾向がある (坊農、2020)。そうしたジャーナルは、通例、刊行までに期間を要するため、新しい知見についての先取権を主張するには、プレプリントの公開によって「幅広い、また、素早い研究成果の共有」をはかることが必須となる (林、2021)。一方、人文系、とくに言語系研究者から見ると、RG の価値は、「幅広い、また、ゆっくり時間をかけた、人目につきにくい研究成果の共有」を可能にしてくれる点に見いだされる。近年、研究者間の競争が激化し、「査読論文しか出さない。他はやっても労力の無駄だ」という意識が研究者の間に広がっているとされるが (坊農、2020)、RG のようなプラットフォームが、査読の有無を問わず、幅広い研究に光を当てる契機になることもある。

#### (5) 新しい研究分野の育成

第二著者の現在の主要な研究フィールドの一つは、学習者コーパス研究である。この分野が誕生したのは2000年前後で、類縁分野であるコーパス言語学と比べても圧倒的に歴史が浅い。この間、当該分野を牽引してきたのはベルギーの研究者だが、彼女は、自身の過去の論文のほぼすべてについて、プレプリントをRGで公開してきた。これにより、各国の学生・大学院生を含め、当該分野に関心を持つすべての研究者が、分野の骨格をなす彼女の研究に無償でアクセスできる環境が整い、彼女が提唱する理論や分析手法の定式化と普及、また、批判的ディスカッションを通じた理論の精緻化がなされた。RGは、当該分野が言語研究の新しい分野として短期間で自立する上で大きな役割を果たしたと言える。現在では、彼女に直接・間接に影響を受ける形で、第二著者を含め、同じ分野の研究者の多くが同様の研究公開を行っている。

### 5. 3 RGの課題

以上で、RGの利点について述べたが、RGなどのプラットフォームには、同時に、留意すべき点も認められる。以下では3つの課題について述べたい。

#### (1) サービスの持続可能性への懸念

RGはこれまで出資者からの支援をもとに急速に拡大してきたが、ビジネスモデルは明瞭ではない。将来的に、研究者と研究機器業者、ないし、求職者と求人機関とのマッチングサービスなどで収益化を行う方向性が示されているが (t.ly/DQ-O)、RG自身が強調するように「RGの価値を作り出しているのは会社ではなく研究を公開している研究者である」以上、研究者の権利や権益を守りつつ、安定した収益化を図ることには困難も予想される。RGが営利団体である限り、RGの提供するサービスの永続性については一定の懸念が残る。

#### (2) 著作権侵害の危険性

有力ジャーナルの多くは有償講読制を採用しており、ジャーナル掲載論文が仮に全く同じ形でRGに公開されたとするなら、ジャーナル出版社の潜在的利益は棄損される。そうした行為は禁止されているわけだが、実際には、プレプリントや著者最終版でなく、「購読費を払わないと読めない出版者版のファイルを不特定多数の研究者と共有する場合があります、著作権上の問題」となる事態も生じている (林、2017)。

2017年10月、ジャーナルの出版元である米国化学会 (ACS) と Elsevier は、有償講読論文がRGに公開されているとして、ドイツの地方裁判所にRGを提訴した。裁判は2022年1月に結審し、RGは公開されるコンテンツに責任を負うという判断が示され、ACSとElsevierの論文の削除が命令された。一方、著作権譲渡手続きの不備などの理由で、出版社による損害賠償請求は却下された。見方によっては勝者のはっきりしない判決で、両者ともに控訴の意向とのことである。

この問題については、コンテンツを公開するプラットフォームの責任を厳しく追及する立場と、オープンサイエンスの公益性を重視する立場があり、今後の裁判の行方は見通しにくい (Kwon, 2022)。しかし、権利関係をめぐる曖昧さ、とくに、研究者が意図せずして著作権侵害を犯してしまうリスクは、RG などの論文共有プラットフォーム利用者全体の懸念事項となりうる。

### (3) 統計データの問題

5.2 節で述べたように、RG には、公開された論文について詳細な統計データを取集し、会員本人にフィードバックする仕組みがある。これは RG の大きな利点であるが、一方で、RG で提供される統計データの妥当性は必ずしも保証されたものではない。たとえば、引用数についても、どのようなクオリティの論文に引用されたかは問われず、自己引用でも増えるという問題がある。また、RG では、「論文を出した数、誰かからの質問に答えた数、followers の数など」に基づき、研究者の総合的なパフォーマンスを数値化する RG Score を算出していたが (坊農, 2020)、計算アルゴリズムはきわめて不透明であった。

もっとも、統計データの質の向上のため、RG が一定の努力をしている事実はあわせて述べておくべきだろう。たとえば、購読数と完全購読数については、登録会員がログイン中に行った購読と、それ以外を区別してカウントするようになっている。また、*h-index* (「*h* 回以上引用された論文が *h* 本以上ある」ことを示す数値。*h-index*=5 ならば「5 回以上引用された論文を 5 本以上有する研究者である」ことを示す) の計算にあたっては、自己引用を含む数字と含まない数値を区別して出力する仕組みが導入されている。さらに、批判の多かった RG Score は 2022 年夏までに廃止された。

RG Score の廃止に伴い、RG は、個人の研究パフォーマンスを多角的に可視化する枠組みと、新しい総合指標値を開発し、2022 年後半より実装が開始された。このことは、今後の研究者評価のあり方を考える上でも参考になる点が多いため、以下、項を改めて簡潔に紹介しておきたい。

### (4) 新しい研究インパクト評価

2022 年夏のシステム変更にあたり、RG は、個々の研究者による「学界へのインパクト (impact in the scientific community)」をより多面的に評価する方針を立てた。そして、これを実装するものとして、新しい情報パネルが設けられた ([t.ly/IWWS](https://t.ly/IWWS))。そこでは、公開されている全論文に基づく総計の購読数、推薦数 (論文を読んだ登録会員が読後に推薦を行った数)、引用数、外部言及数 (ブログや SNS など、RG の外部で論文への言及・リンクがあった数) が一括表示される。あわせて、RG Score に代わるものとして、Research interest score (RIS) という指標値が新設された。

計算の仕組みがブラックボックスであった RG Score の反省に基づき、RIS は計算のアルゴリズムが公開されている ([t.ly/Z8Jr](https://t.ly/Z8Jr))。RIS の背景にあるのは、研究者が他者の研究を受容・評価するという行為を、「A: ある研究を見つけてアクセスする」、「B: 興味を持てば全文を読む」、「C: 内容を評価すれば当該論文を推薦する」、「D: さらに価値があると思えば自身の研究で引用する」の 4 段階に区分してとらえる考え方である。この発想に沿って、A を部分購読数 (アブストラクトや図表などを見た回数) (ウェイト: 5%)、B を完全購読数 (15%)、C を推薦数 (25%)、D を引用数 (50%) で計量化し、それぞれに所定のウェイトをかけて合算したのが RIS である。RIS は論文別に計算され、その積み上げで、研究者ごとの指標値が決まる。

指標の信頼性を向上させるため、非登録会員による購読、登録会員が非ログイン中に行った購読、ボットやクローラーによる機械的なアクセス、同一研究者による 1 週間以内の反復アクセス、著者自身の自己購読や自己引用などはすべて機械的に排除され、計算がなされる。また、分野の

違いや研究歴の長短の影響を考慮するため、分野別、研究歴別の RIS ランクも表示される。

もちろん、機械的な計算である以上、RIS にも改善点が残されていることは確かであるが、研究のインパクトをできるだけ多面的に計ろうとする方向性は注目に値する。従来、研究評価においては、引用数が絶対の基準であったが、引用数は、分野の影響を強く受けるほか、レビュー論文のほうが高くなりやすいといった課題もある。この点に関して、RG は、「引用数は古くから研究評価の指標とされてきたが、実際には、引用されるまでに、数カ月ないし数年を要することもある。その間にも、多くの研究者が当該論文を読み、そこから学びを得ている。ゆえに、引用数だけを見ることは研究のインパクトを過小評価しかねず、講読数・推薦数・引用数を複合的に見て評価すべきである」と述べている（第二著者による抄訳）。RG の見解は、学術界における研究評価・研究者評価にも一石を投じるものと言えよう。

#### 5. 4 大学レポジトリと RG

目的と機能において、RG と重なる部分が多いのが、大学の運営する機関レポジトリである。そのため、レポジトリを運営する大学図書館から見た場合、RG に対する評価は複雑なものとなりがちである。構成員の研究がより広く認知されるというメリットの一方で、構成員が著作権のトラブルに巻き込まれる危険性が生じたり、RG にのみ論文が投稿されることでレポジトリの相対的な価値が低下したりするデメリットが予想されるからである。

海外の大学の図書館では、教員や学生向けに RG の紹介を行っているところも少なくない。オランダの Erasmus University Rotterdam は、研究者が「学術的な可視性と注目度を高める」手段として RG などを紹介しており、会員登録方法の詳細な解説のほか、RG での論文公開が Google Scholar のインデックス対象になることも示されている ([t.ly/Gh9M](https://t.ly/Gh9M))。また、米国の Virginia Tech では、RG の主なサービス内容とともに、RG が営利団体であることや、著作権問題への懸念があることも示され、代替となる大学独自のレポジトリも紹介されている ([t.ly/lzVy](https://t.ly/lzVy))。カナダの University of Saskatchewan では、RG の運営者、会員登録方法、RG が採用する各種統計値の解説、会員プロフィールページの作成例など詳細に示されるほか、研究者 SNS を使用する際の留意点も丁寧に解説されている ([t.ly/3L7W](https://t.ly/3L7W))。

一方、国内の大学図書館では、研究者向けの RG の直接的な紹介はほとんどないようである。関西圏の 3 つの国立大学図書館の公式サイトを調査したところ、A 大学では RG に関して 2 件のヒットがあり、1 件は学生向けの「論文の探し方ガイド」での言及、他の 1 件は教員向けイベント告知の中での言及であった。B 大学ではヒットは 1 件で、学生向けの「テーマ別調べ方ガイド」内での RG 公開論文へのリンクであった。C 大学では 23 件のヒットがあったが、大半は RG 論文を参考文献に含めたレポジトリ論文へのリンクで、残りは、図書館職員向け研修会資料での言及、また、レポジトリの広報誌上で行われた教員インタビュー内での言及であった。

Virginia Tech の解説にもあるように、研究者から見た場合、大学が運営する機関レポジトリは、RG などの代替となりうる存在である。レポジトリは、サービスの継続性が担保されているほか、プレプリントの公開にあたっては、図書館員が著作権関係を代行調査し、問題がないことを確認して公開する手順がふまれるため (菅原、2020)、出版社の著作権を侵害してしまう恐れもほぼ皆無である。レポジトリであれば、プレプリントだけでなく、著者最終版や出版最終版の公開を認めるジャーナルもある。ただ、「幅広い読者へのリーチ」という点に関して言えば、現状では、レポジトリと RG の差は大きい。第二著者の論文の一部は、RG と所属先大学の機関レポジトリで

重複公開されているが、RG で最もよく読まれている論文を例にして言うと、レポジトリからのダウンロード数はRG の約4分の1にとどまっている。

もともと、今後、知名度が上がっていけば、機関レポジトリは幅広い研究者にとってさらに魅力的な研究成果の公開先となるだろう。[河合 \(2020\)](#) は、すでに体制の整っている機関レポジトリが、現在の中心コンテンツである学内紀要論文や博士論文だけでなく、「プレプリント公開や、その発展版である、より自由度の高いコンテンツ公開」に踏み切ることで、研究者が「レポジトリにメリットを感じていない現状」は変化するだろうと述べている。

近い将来、研究者は、幅広いプラットフォームの中から、研究の内容や公開の目的に応じて適切な公開先を一つないし複数選ぶことになっていくと思われるが、その選択は容易ではない。とくに、若手研究者に対するこうした面での指導はほとんどなされていないため、本稿4節で述べたような取り組みの重要性はさらに高まっていくものと思われる。

## 6. まとめ

以上、本稿においては、言語系研究におけるオープンサイエンスの動向を2つのプラットフォームに着目して概観した。第2~4節においては、Jxiv の概要を示すとともに、第一著者の体験をふまえ、Jxiv に投稿前プレプリントを公開する意義と、研究者教育における活用の方策について検討した。第5節においては、第二著者の経験をふまえ、RG に刊行済み論文のプレプリントを投稿する利点や問題点、また、機関レポジトリとの関係について考察した。近年のオープンサイエンスの潮流はすべての分野の研究者に大きな変化を迫るが、変化を前向きに受け止めることで、言語系研究をさらに活性化していくことが可能であると思われる。

## 注・謝辞

本稿の内容の一部は、2022年6月21日に開かれた国立国語研究所「第241回NINJALサロン」における第一著者の口頭発表「昨今の学術研究環境とJxivの動向」に基づく。この発表において、鳥取大学医学部専任講師の柴田敏史氏よりコメントを、岩崎拓也氏と宮川創氏からは指定討論者として助言をいただいた。また、本稿5節の執筆にあたっては神戸大学附属図書館電子図書館担当者より助言を得た。加えて、相澤正夫氏には、完成原稿をお読みいただき、貴重なコメントをいただいた。各位に感謝申し上げます。ただし、残された問題はすべて著者両名の責任である。

## 参考文献

- Kwon, D. (2022). ResearchGate dealt a blow in copyright lawsuit. *Nature*, 603, 375-376. <https://www.nature.com/articles/d41586-022-00513-9>
- Singh Chawla D. (2022). Japan launches preprint server - but will scientists use it? *Nature*, 10.1038/d41586-022-01359-x. Advance online publication. <https://doi.org/10.1038/d41586-022-01359-x>
- 有田正規 (2021) 『学術出版の来た道』岩波書店。
- 尾城孝一 (2021) 「学術論文のバージョンとプレプリントをめぐる動向」情報科学技術協会 INFOSTA OUG ライフサイエンス分科会 2021年11月度(第388回)資料. [https://www.infosta.or.jp/wp/wp-content/uploads/2021/12/OUG\\_LIFE\\_202111handout.pdf](https://www.infosta.or.jp/wp/wp-content/uploads/2021/12/OUG_LIFE_202111handout.pdf)
- 河合将志 (2020) 「機関レポジトリによるプレプリント公開」国立情報学研究所第2回 SPARC Japan セミナー2020 発表資料. [https://www.nii.ac.jp/sparc/event/2020/pdf/20201218\\_doc2.pdf](https://www.nii.ac.jp/sparc/event/2020/pdf/20201218_doc2.pdf)

- 柴田敏史 (2020) 「学会の持続可能性を模索する」に対するコメント」第 241 回 NINJAL サロン「昨今の学術研究環境と Jxiv プレプリントの動向」発表資料 (発表者: 横山詔一)  
[https://researchmap.jp/multidatabases/multidatabase\\_contents/detail/230097/5c7bcee6ffdc17809a5c116cec92761d?frame\\_id=835325](https://researchmap.jp/multidatabases/multidatabase_contents/detail/230097/5c7bcee6ffdc17809a5c116cec92761d?frame_id=835325)
- 菅原光 (2020) 「オープンアクセス・機関リポジトリ」国立大学図書館協会東京地区協会・関東甲信越地区協会合同フレッシュパーソンセミナー講演資料.  
[https://www.janul.jp/sites/default/files/2020-11/r2\\_fps\\_sugawara.pdf](https://www.janul.jp/sites/default/files/2020-11/r2_fps_sugawara.pdf)
- 田辺浩介・谷藤幹子 (2021) 「オープンサイエンス時代の研究者プロフィールサービス—研究活動の可視化のためにやるべきことはなにか?—」『情報の科学と技術』71(5), 200-205.  
[https://www.jstage.jst.go.jp/article/jkg/71/5/71\\_200/\\_pdf-char/ja](https://www.jstage.jst.go.jp/article/jkg/71/5/71_200/_pdf-char/ja)
- 日本学術会議オープンサイエンスの深化と推進に関する検討委員会 (2020) 『提言 オープンサイエンスの深化と推進に向けて』 <https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t291-1.pdf>
- 林和弘 (2017) 「学術情報流通のオープン化がもたらすオープンサイエンスに向けた成果公開プロセスと共有の変革」『STI Horizon』3(3), 35-39. <https://www.nistep.go.jp/wp/wp-content/uploads/NISTEP-STIH3-3-00092.pdf>
- 林和弘 (2021) 「論説: 本格的に変わり始めた科学と社会」『化学と工業』74(10), 719-720.  
<https://www.chemistry.or.jp/opinion/ronsetsu2110.pdf>
- 坂東慶太 (2018) 「動向レビュー: 研究者 SNS とそこに収録された文献の利用」『情報の科学と技術』68(4), 189-195. [https://www.jstage.jst.go.jp/article/jkg/68/4/68\\_189/\\_pdf-char/ja](https://www.jstage.jst.go.jp/article/jkg/68/4/68_189/_pdf-char/ja)
- 坊農秀雅 (2020) 「生命科学研究におけるプレプリントや SNS 活用の現状と課題」国立情報学研究所 第 2 回 SPARC Japan セミナー 2020 発表資料.  
[https://www.nii.ac.jp/sparc/event/2020/pdf/20201218\\_doc5.pdf](https://www.nii.ac.jp/sparc/event/2020/pdf/20201218_doc5.pdf)
- 横山詔一・前田忠彦・高田智和・相澤正夫・野山 広・福永由佳・朝日祥之・久野雅樹 (2022) 「日本人の読み書き能力 1948 年調査における非識字率と生年の関係」Jxiv プレプリント:  
<https://doi.org/10.51094/jxiv.73>

## Research and Teaching in Language Studies in an Age of Open Science: Use of Preprint Servers

### Authors:

YOKOYAMA Shoichi, Research Department, National Institute for Japanese Language and Linguistics  
ISHIKAWA Shin'ichiro, IPHE/Graduate School of Intercultural Studies/CMDS, Kobe University

### Abstract:

This article focuses on the growing significance of preprint sharing for research and teaching in language studies in an age of open science. The authors firstly touch upon the background of the establishment of Jxiv, a new preprint server (Section 2), discuss ethical aspects of sharing preprints (Section 3), and introduce how to teach research ethics to graduate students (Section 4). Then, they discuss the merits and drawbacks of using ResearchGate by comparing it to the institutional repositories (Section 5).

Keywords: language studies, open science, preprint servers, Jxiv, creative commons (CC) licenses, teaching for graduate students, researcher SNS, ResearchGate, institutional repositories

## 倫理チェックリスト

以下 Q1 から Q13 まで、末尾に同意文書の書式も添付

- 利益相反（Conflict of Interest: COI）の開示 → Q12 で記述
- 著者の貢献（Author Contributions）の記載（著者が複数の場合） → Q9-2 で記述

Q1. 研究を行うにあたり、所属または関連機関の倫理委員会の承認を得ましたか。

A1. いいえ。

理由：本研究は文献資料と著者自身の体験にもとづくものであり、人間や動物を対象にした研究ではない。

Q2. 実験や調査に先立ち研究参加者からインフォームド・コンセントを得ましたか。

A2. いいえ。

理由：上記 A1.と同じ。

Q3. やむをえずインフォームド・コンセントが得られない場合は、代替となる手段をとりましたか。（親や責任者による承諾を得るなど）

A3. いいえ。

理由：上記 A1.と同じ。

Q4. 実験や調査においては、研究参加者や動物に負荷やリスクはありませんでしたか。

A4. 該当しない（人間や動物を対象にした研究ではない）。

Q5. 実験や調査にデセプションがありましたか。

A5. いいえ。

Q6. 動物実験においては、必要最小限の個体数で実験しましたか。

A6. 該当しない（動物実験は実施せず）。

Q7. プライバシーは保障されていますか。

A7. はい。

Q8. 論文は著者自身によるオリジナルの論文ですか。

A8. はい。

Q9-1. 著者が連名である場合、連名者全員から投稿の承諾を得ていますか。

A9-1. はい。さらに、クリエイティブ・コモンズ・ライセンスの表示について「CC BY」とすることについても連名者全員の賛同・承諾を得ている。この点の詳細は Q13 も参照されたい。

Q9-2. 著者名の順序は貢献度を適切に反映していますか。

A9-2. はい。具体的な貢献度は連名著者順に次のとおりである。

横山詔一：論文の構想を提案，論文執筆，投稿の統括

石川慎一郎：論文の構想について確認・吟味，論文執筆、論文全体のとりまとめ

Q10-1. 他者が作成した材料やプログラムを用いたり，図表や本文を引用したりした場合，その出典は示されていますか。

A10-1. はい。

Q10-2. 原著者からの承諾を得ていますか。

A10-2. 該当しない（現著者からの承諾が必要な引用はない）。

Q11. 不適切，あるいは差別的な用語や表現がないかチェックしましたか。

A11. はい。

Q12. 企業などと共同研究を実施，あるいは企業などからの助成を受けましたか。（利益相反（COI：Conflict of Interest）について，研究の公正性，信頼性を確保するためには，利害関係が想定される企業などのかかわりについて，適切に対応する必要があります。）

A12. いいえ。

Q13. 共著者との共同研究である場合，各著者はその公表の仕方についても相互に十分な説明のもとで同意をしていますか。（クリエイティブ・コモンズ・ライセンスの表示については，全ての著者がその具体的に意味するところを十分に承知している必要があります。）

A13. はい。同意文書に残すかたちで十分な確認をしている。

参考：同意文書の書式

Jxiv 投稿にあたっての同意について

(筆頭著者氏名) 様

私は次に示す論文の共著者として、Jxiv への投稿、及びクリエイティブ・コモンズ・ライセンスの表示について十分に理解したうえで、下記の3点すべてに同意します。

論文表題：(○○…)

共著者名：(筆頭著者・共著者 A・B・C…)

記

1. Jxiv に上記の共著者順により投稿すること
2. 投稿論文のクリエイティブ・コモンズ・ライセンスの表示を「CC BY」とすること
3. その後に学術雑誌等に投稿する場合も、クリエイティブ・コモンズ・ライセンスの表示を「CC BY」とすること

2022 年\*月\*日  
(共著者自署)