

# 日本の研究者のプレプリントに 対する捉え方と今後の展望

大森 悠生(おおもり ゆうき)\* 池内 有為, 逸村 裕, 林 和弘

\*筑波大学大学院 情報学学位プログラム, 科学技術・学術政策研究所

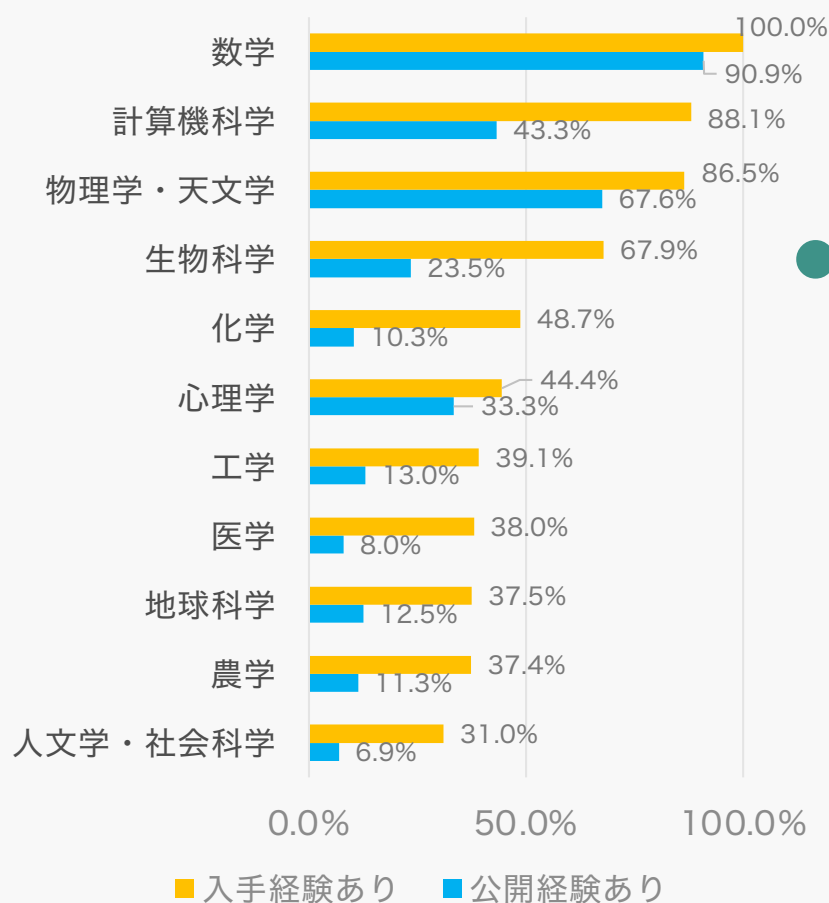
# 目次

1. 研究背景
2. 先行研究
3. リサーチクエスチョン
4. 調査方法
5. 結果
6. 考察
7. 結論

## プレプリントとは

- 学術雑誌へ投稿する予定の論文の草稿
  - プレプリントサーバにおいて共有されることが多い
    - プレプリントサーバ: プレプリントを共有するインターネット上のサーバ
      - 例: arXiv(物理学など)・bioRxiv(生物学など)・medRxiv(医学など)
- 特徴
  - 投稿後、短期間にサーバで公開される
    - 雑誌論文は投稿後、査読などで公開までに時間がかかる
  - 誰でも無償で利用が可能
    - 無料・無制限に公開されるため、Green OAを実現する手段の一つとしても注目
  - 査読が存在しない
    - 最低限のチェック・スクリーニングは行われているが、新規性等は確認されない

## 日本の研究者のプレプリント入手・公開状況



### ● 一定の研究者は入手・公開経験あり

● 入手経験あり：5割程度

● 公開経験あり：2割程度

→数学, 物理学・天文学, 計算機科学等で高い

図1, 日本の研究者のプレプリント入手・公開状況<sup>(1)</sup>(n=1,448)

## COVID-19によるプレプリント公開の増加

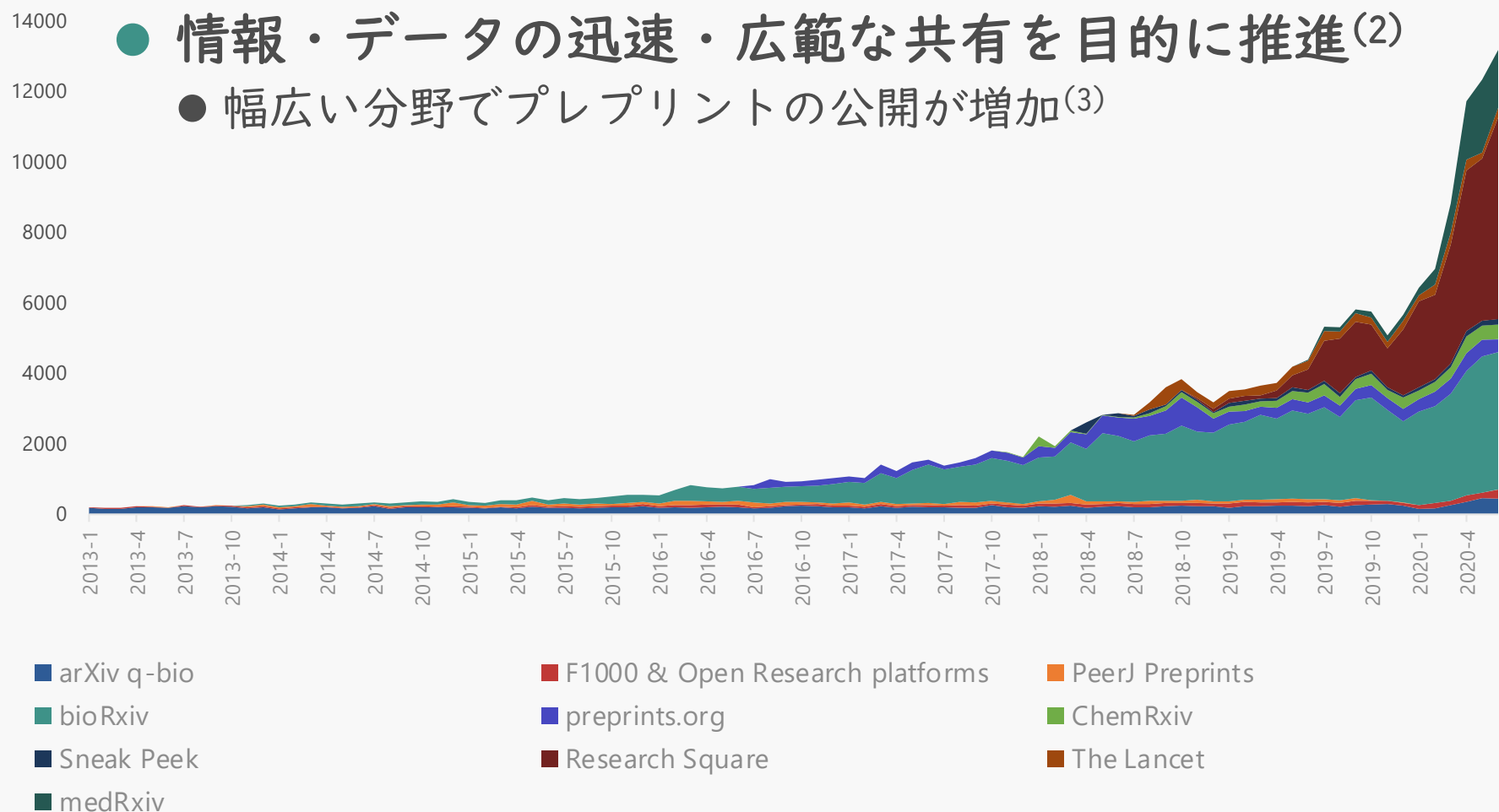


図2, プレプリントサーバごとに投稿されたプレプリントの数<sup>(4)</sup>

## プレプリントに対する注目

- プレプリントに言及するSNS上の投稿<sup>(5)</sup>
  - プレプリントを投稿したところ,翌日には注目が集まる<sup>(6)</sup>
  - COVID-19に関するプレプリントへの多数の言及
- プレプリントを根拠とした報道<sup>(7)</sup>
  - プレプリントを根拠とした報道を行う
  - プレプリントの信頼性などの課題についても触れている

→研究者でなくともプレプリントに触れる機会が増加

## 池内, 林(2021)<sup>(8)</sup>

### ● 目的

- 日本の研究者のプレプリントの利活用状況と認識を明らかにする

### ● 方法

- 日本の研究者(約1,500)名に対する質問紙調査(2020年8月～9月実施)

### ● 結果

- 分野差はあるが約半数の研究者が入手・約2割が公開経験があった
- 成果の広報や先取性の確保を目的に公開している研究者が多い
- 最初に査読誌へ投稿するために公開していない研究者が多い

### ● 考察

- プレプリントは全ての分野に浸透しているとは言えない

## Fraserら(2021)<sup>(9)</sup>

### ● 目的

- COVID-19に関連するプレプリントの利用・伝播状況を調査する

### ● 方法

- bioRxivとmedRxivに投稿されたプレプリントの利用状況を収集

### ● 結果

- COVID-19関連プレプリントの投稿は他の感染症と比較して多い
- DL数や引用数なども他トピックのプレプリントと比較して多い

### ● 考察

- 急速なCOVID-19の感染拡大状況下においてプレプリントは有用
- 質保証のないものが報道や政策に用いられている点は課題である



## プレプリントの現況

- プレプリントの利用が推進されている
  - 利用・公開する研究者は一定以上存在する
  - COVID-19などのトピックにおいては研究者外からの利用も多い
- しかし
  - 実際に利用・公開する研究者がプレプリント自体をどのように捉えているかについてはあまり分析がなされていない

## リサーチクエスチョン

1. 日本の研究者はプレプリントをどのように捉えているのか
  - プレプリントに存在する利点と課題
2. 日本の研究者はプレプリントの展望をどのように考えているのか
  - 今後もプレプリントの利用・公開は進むのか
  - プレプリントの解決すべき課題は何か

## ● 調査方法

### ● Web質問紙調査(Cuenote)

- 「論文のオープンアクセスとプレプリントに関する実態調査」

## ● 分析方法

- 選択式回答の集計結果<sup>(10)</sup>のうち、プレプリントに関する回答の再分析
- 自由回答のテキストマイニング

## ● 調査対象

- 科学技術専門家ネットワークに登録されている研究者
  - 科学技術・学術政策研究所の科学技術予測センターが運営する研究者等の2,000人規模のネットワーク

## ● 調査依頼送付数・回答率

- 1,671名に送付・1,173名が回答(回答率70.2%)

## ● 調査期間

- 2022年7月19日～8月22日

## 回答者の分野・年代・所属



図3, 回答者の分野 (n=1,174)

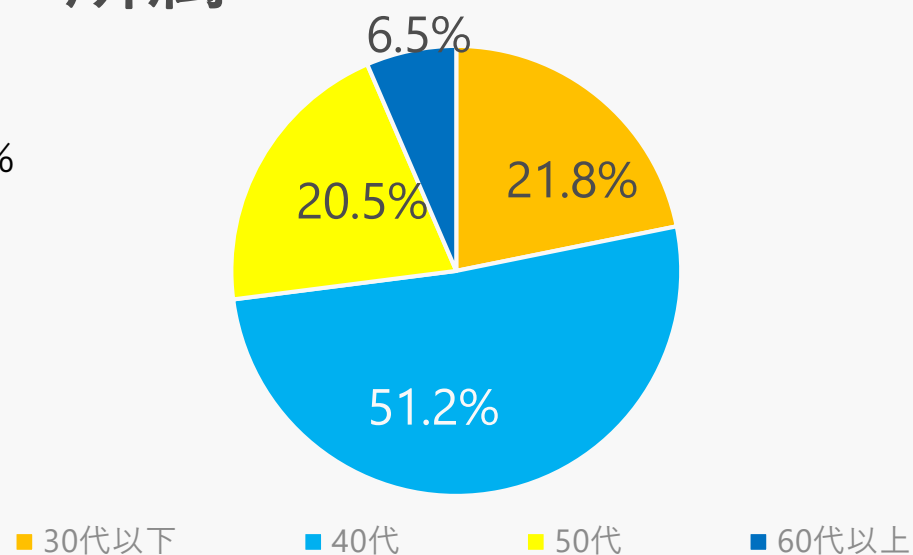


図4, 回答者の年代 (n=1,174)

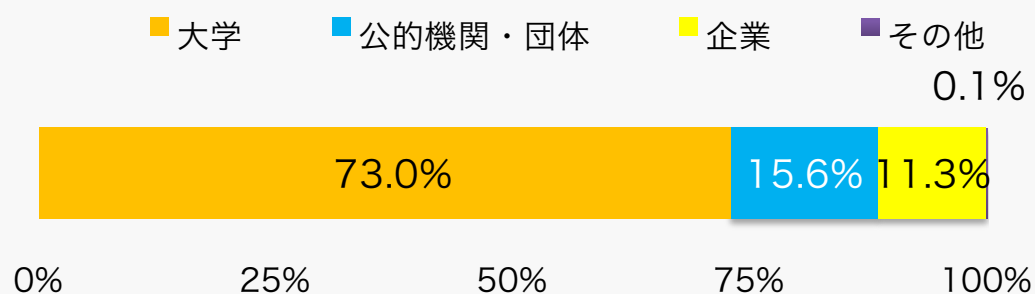
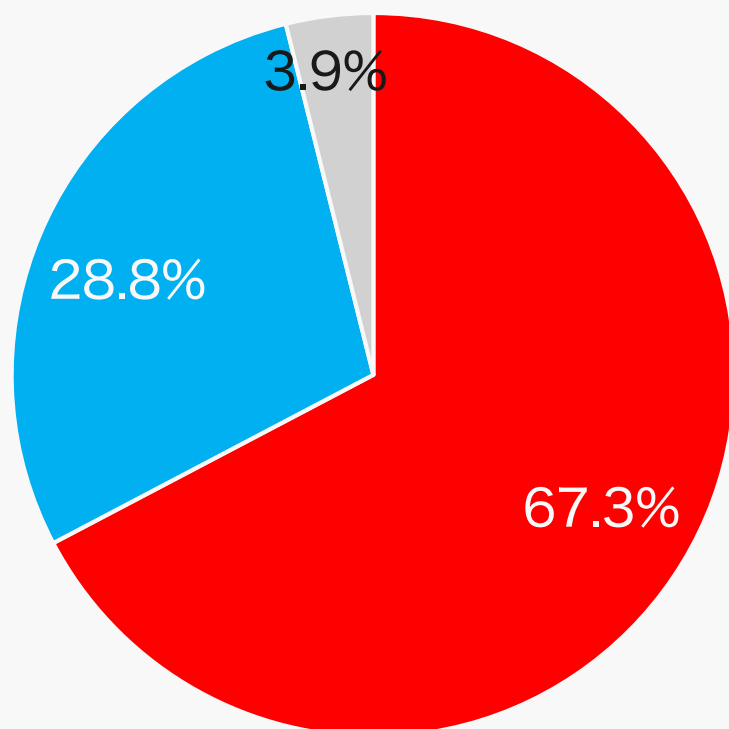


図5, 回答者の所属 (n=1,174)

## プレプリントの利用・公開経験

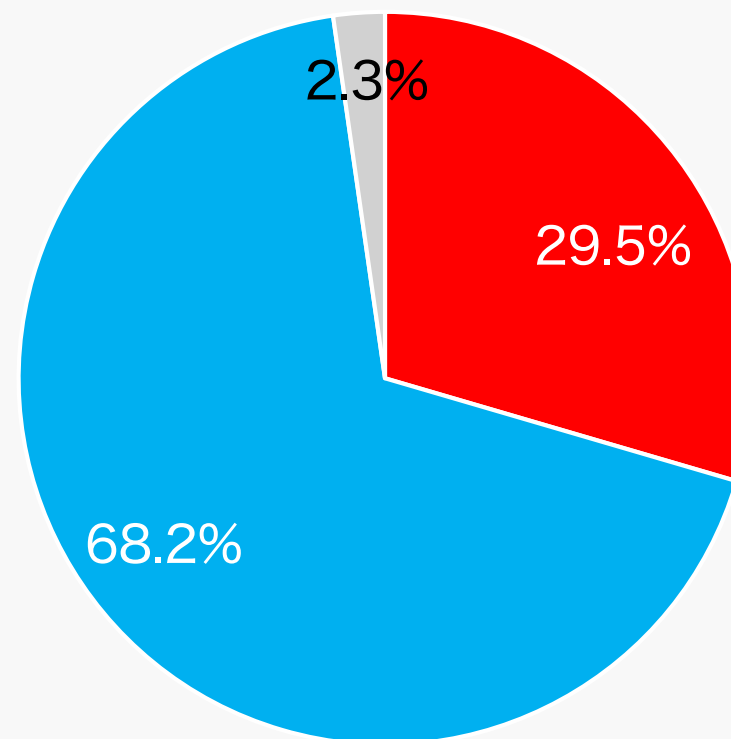
利用経験の有無



■ 1. はい ■ 2. いいえ ■ 3. わからない

図6, プレプリントの利用経験(n=1,104)

公開経験の有無



■ 1. はい ■ 2. いいえ ■ 3. わからない

図7, プレプリントの公開経験(n=1,104)

## 分野ごとのプレプリント利用経験

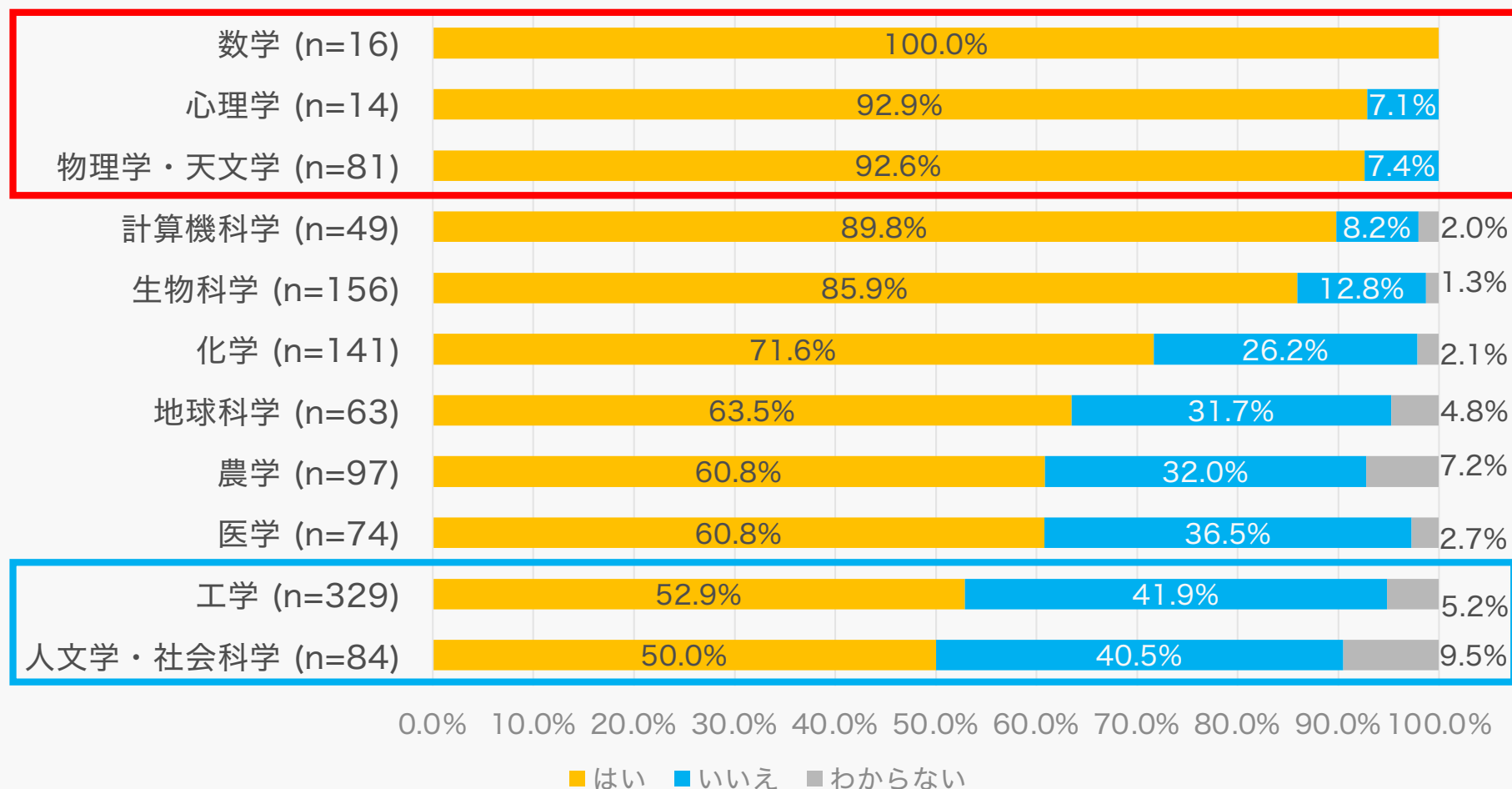


図8, プレプリント利用経験(分野ごと, n=1,104)

## 分野ごとのプレプリントの公開経験

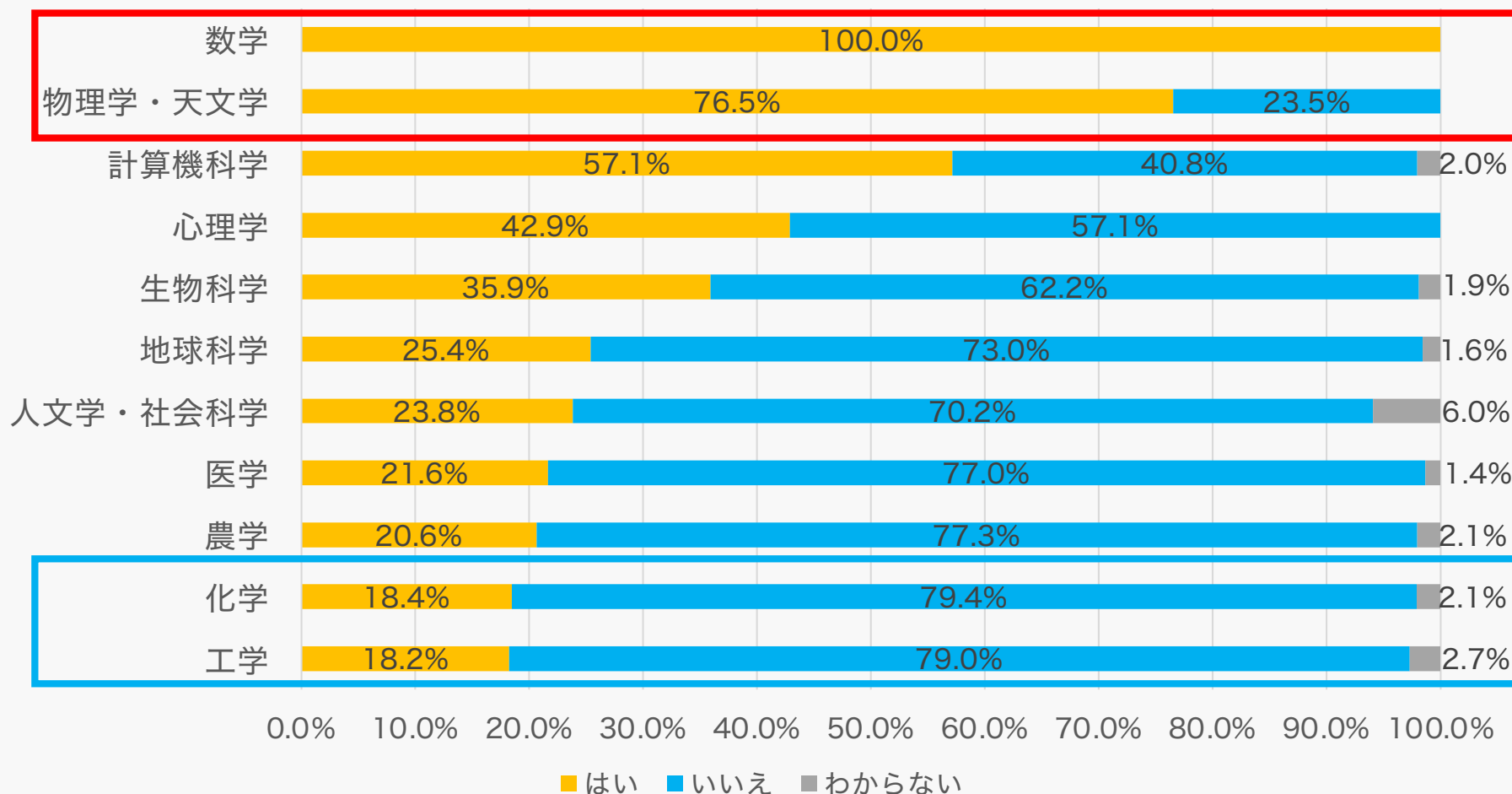


図9, プレプリント公開経験(分野ごと, n=1,104)

## プレプリントを公開した理由

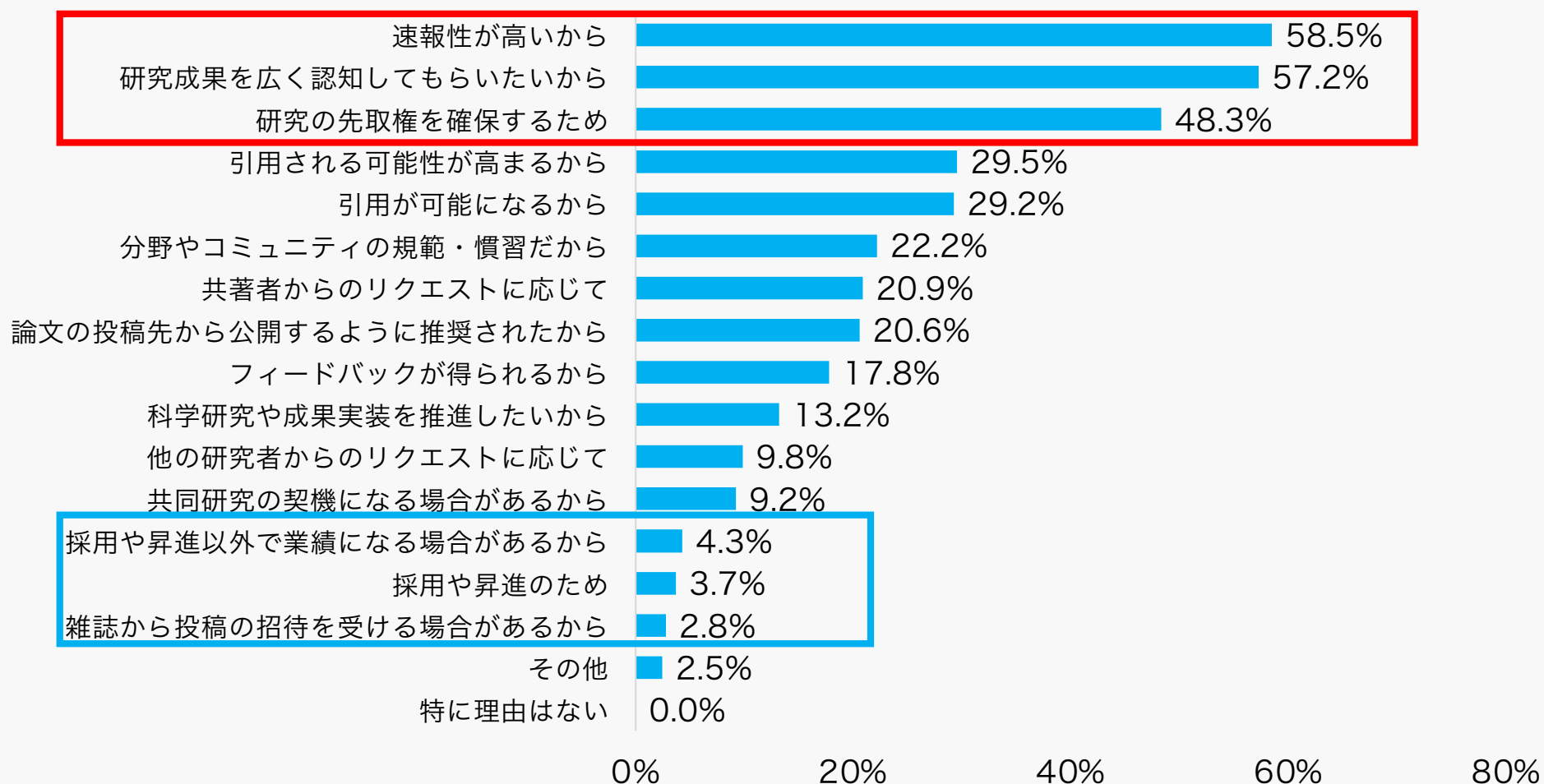


図10, プレプリントを公開した理由(n=326)



## プレプリントを公開していない理由

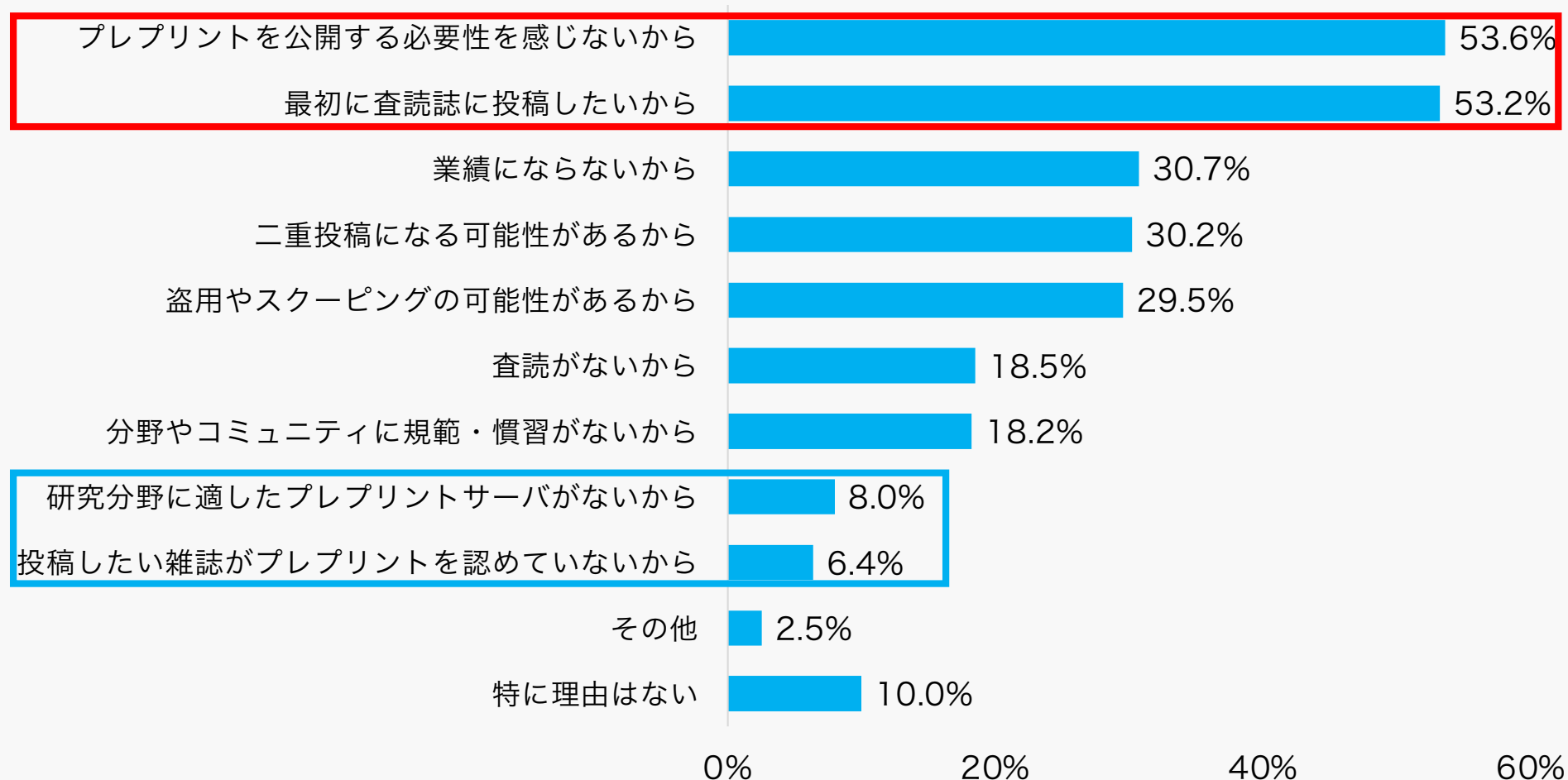


図10, プレプリントを公開していない理由(n=752)

## 公開していない理由が改善されたら公開するか

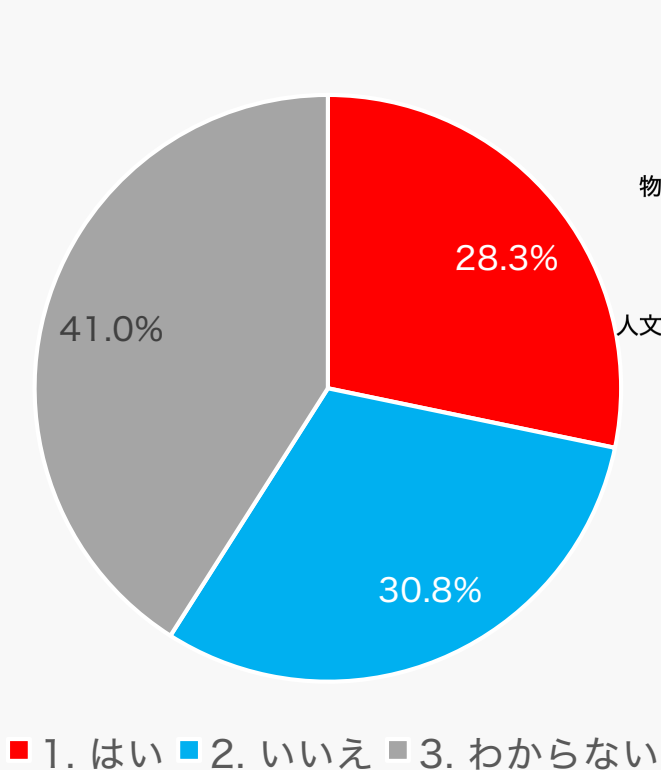


図11, プレプリントを公開していない理由が改善されたら公開するか (n=676)

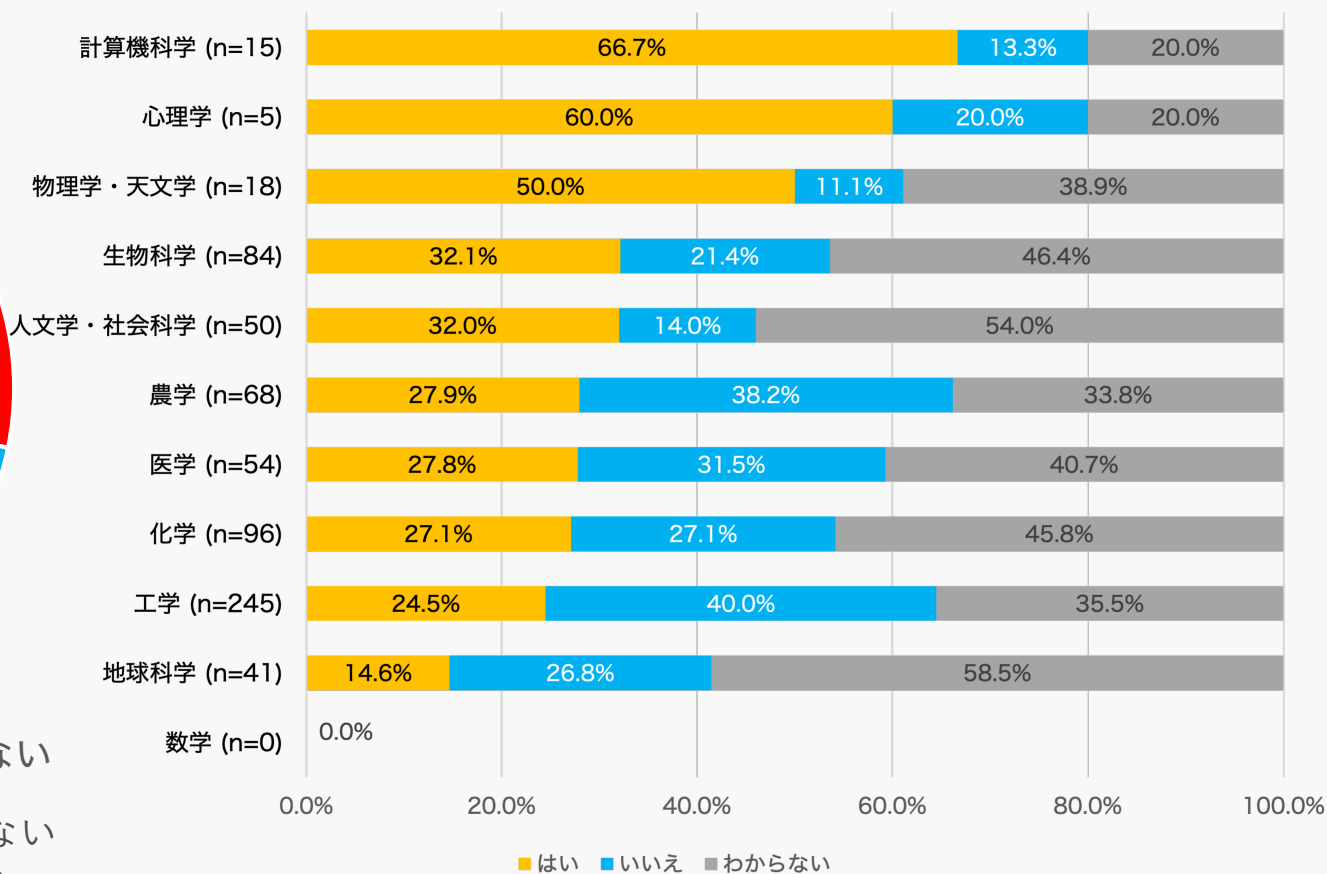
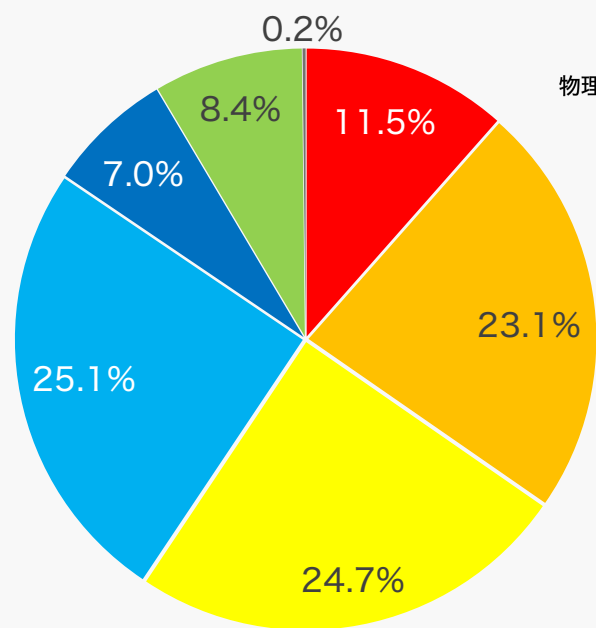


図12, プレプリントを公開していない理由が改善されたら公開するか(分野ごと, n=326)

## 今後、自らの分野でプレプリントの利用が進むと思うか



- 既に十分利用されていると思う
- 進むと思う
- やや進むと思う
- あまり進まないと思う
- 進まないと思う
- わからない
- その他

図13, 今後自らの分野でプレプリントの利用が進むと思うか(n=1,172)

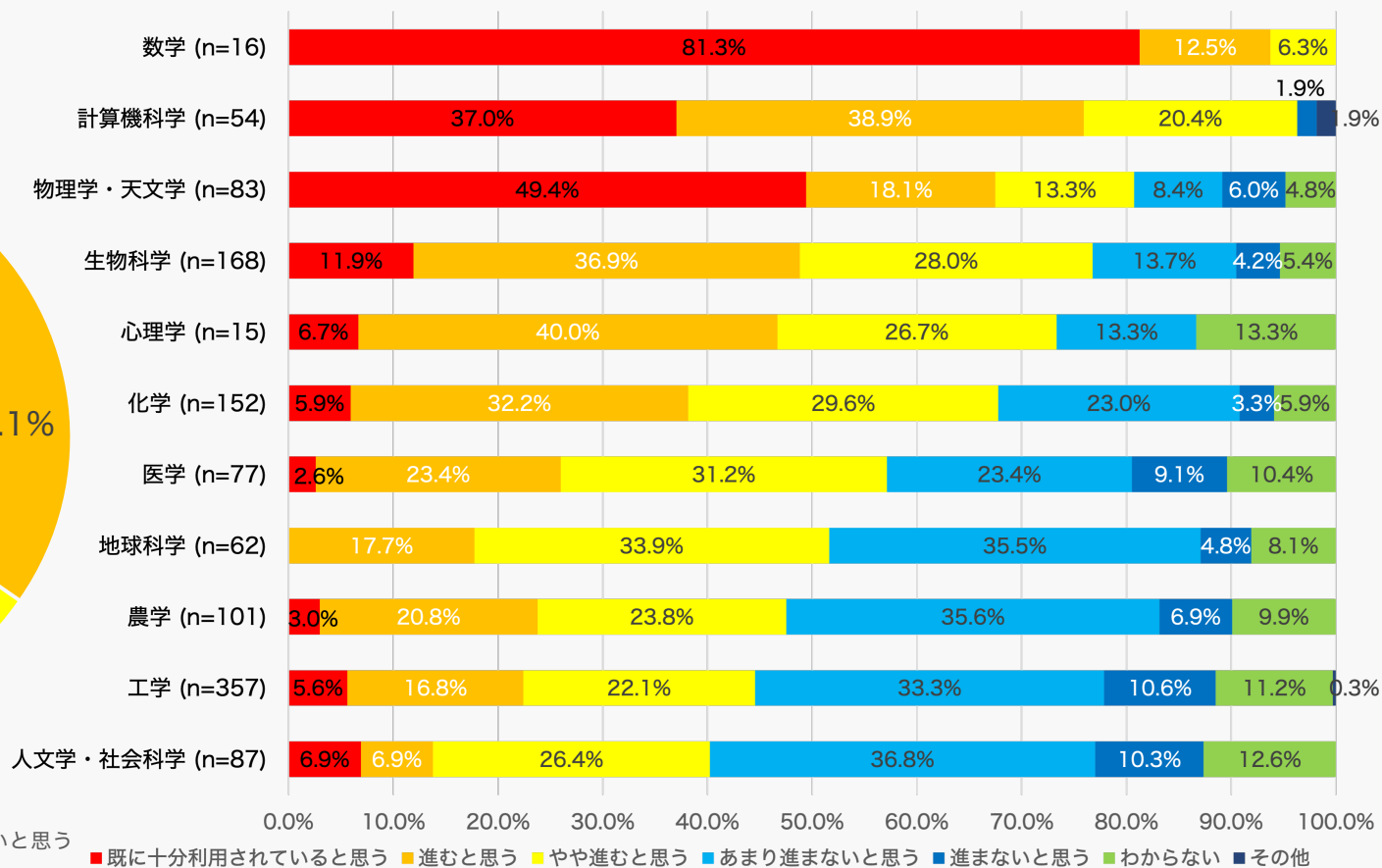


図14, 今後自らの分野でプレプリントの利用が進むと思うか (分野ごと, n=1,172)

# 今後プレプリントを業績として認めるべきか

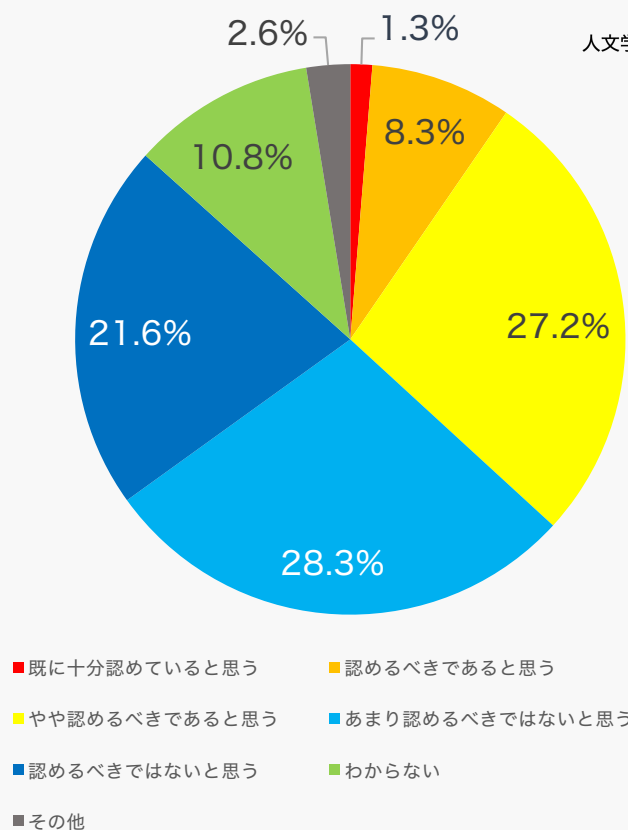


図15, 今後プレプリントを業績として認めるべきと思うか(n=1,172)

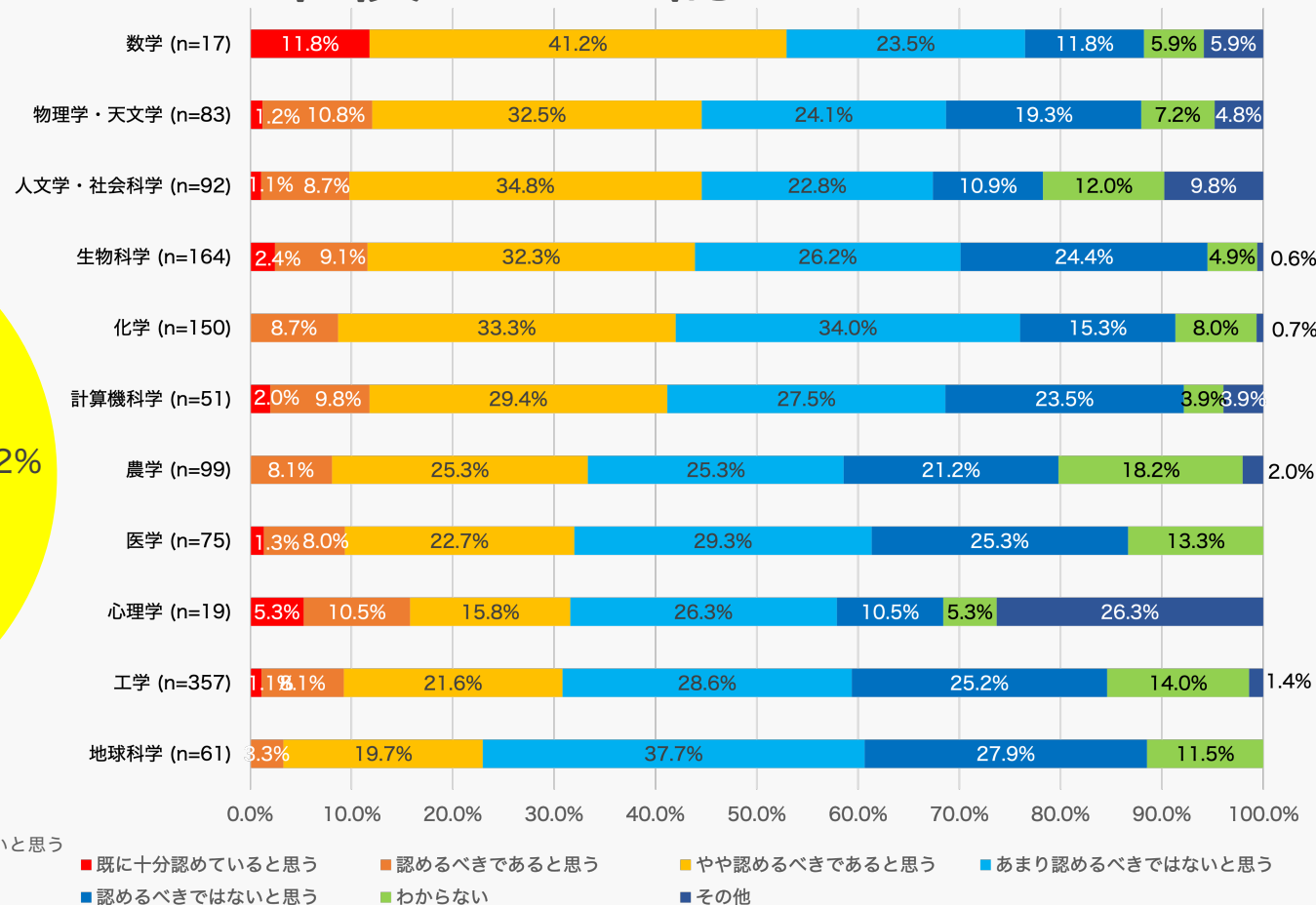


図16, 今後プレプリントを業績として認めるべきか (分野ごと, n=1,172)

## 今後プレプリントを業績として認めるべきか

### ● その他(自由回答)の内訳

- “一定の条件下・基準に基づき認めるべき” という回答が多い

表1, その他(自由回答)の分類

回答内容	回答数
内容や査読の有無による	10
分野による	5
独自の評価基準を設けるべき	5
引用数等の基準による	4
雑誌への投稿とセットで評価	3
別の形態と同等の評価	2
認めるべきだとは思いますが信頼性が欠如	1

## プレプリントの信頼性の判断基準

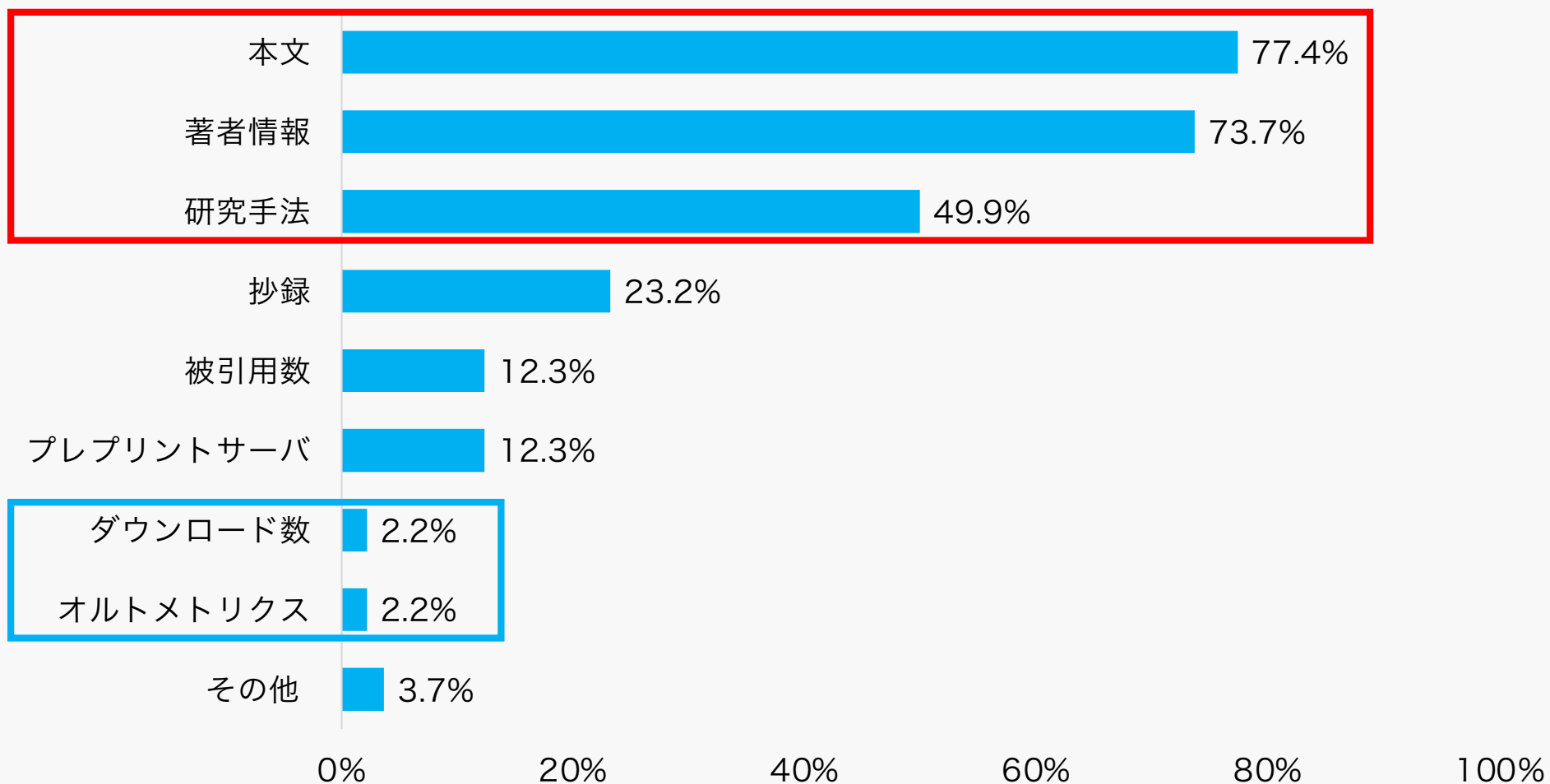


図17, プレプリントの信頼性を判断する基準(分野ごと, n=737)

## 選択式回答の集計結果

### 1. 全分野でプレプリント入手・公開経験のある研究者が存在

### 2. プレプリントの公開・未公開理由

- 公開理由：速報性・研究成果の広報・先取性の確保など
- 未公開理由：必要性を感じない・最初に査読誌へ投稿したいなど
  - 全ての分野に未公開理由が解消されたら公開する研究者が存在する

### 3. 今後のプレプリントの利用

- 全ての分野に今後利用が進む・業績と認めるべきとする研究者が存在
  - 業績として認めるべきか否かは意見が分かれる

### 4. プレプリントの信頼性の判断基準

- 本文・著者情報・研究手法などを基準にする研究者が多い

## テキストマイニング

### ● 質問紙調査における自由回答を対象

#### ● 質問文

- 「論文のオープンアクセスやプレプリントについて、評価、発展性、問題点など、ご自由にお書き下さい。」

#### ● 量的な回答では捉えきれない研究者の考えを明らかに

- KH Coder<sup>(11)</sup>を用い、階層的クラスター分析と共起ネットワークを作成



## 階層的クラスタ分析



図18, 階層的クラスタ分析の結果から一部抜粋

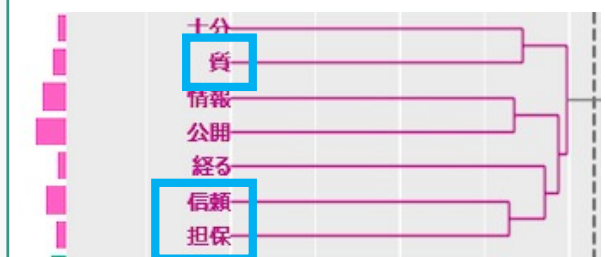


図19, 階層的クラスタ分析の結果から一部抜粋

### ● 階層的クラスタ分析

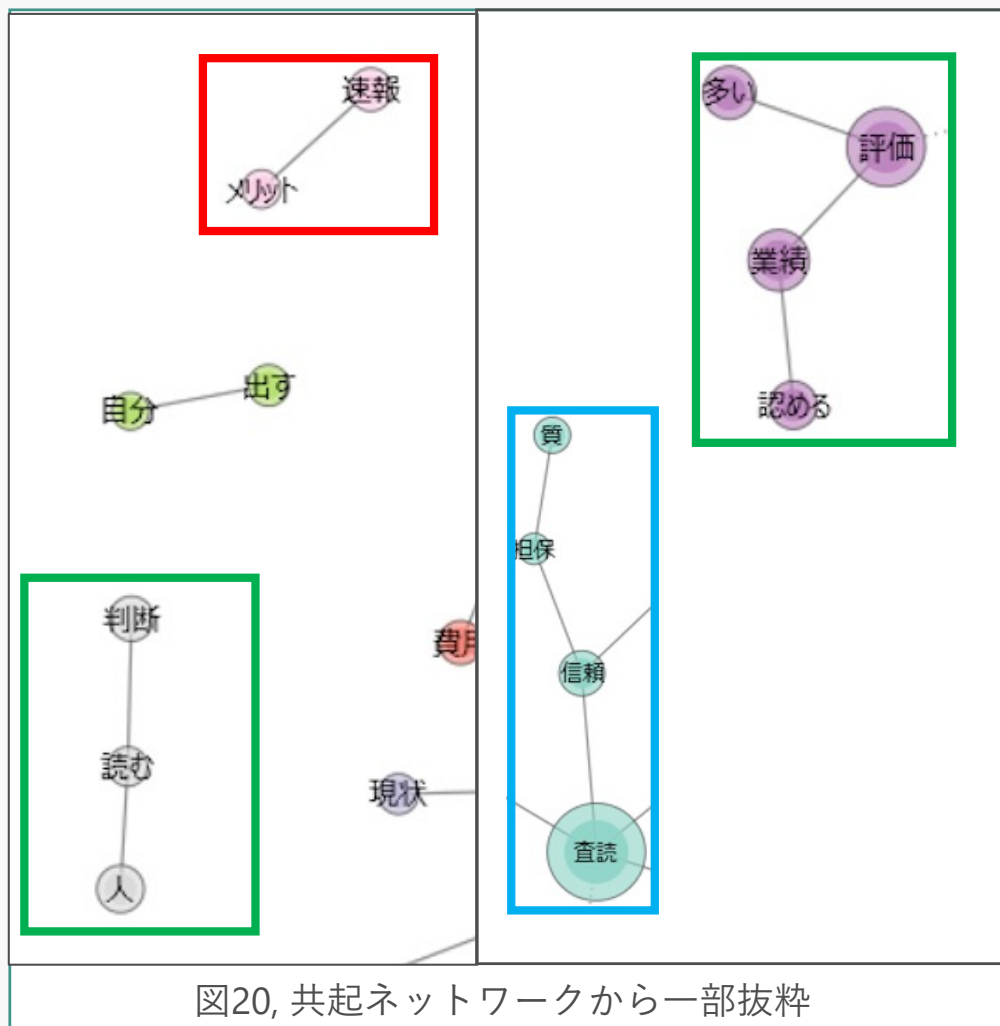
- Ward法を使用

### ● 結果

- 「メリット」「速報」が同じクラスター

! 「質」「信頼」「担保」が同じクラスター  
回答例：“速報性については有用であるものの、信頼性についてはどこまでのものか”

## 共起ネットワーク



- 共起ネットワーク

- Jaccard係数: 0.15

- 結果

- 「速報」 「メリット」

- ! 「信頼」 「データ」 「質」 「担保」

- △ 「読む」 「人」 「判断」

回答例: “読む側が個々に信頼性を検証しなければならず”

- △ 「評価」 「業績」 「認める」

回答例: “査読がついていないので成果として認めるべきかは難しいところ”

“プレプリントは条件付きで認めるのが妥当”

“業績として認めると業績の水増しなど不正の温床になる可能性がある”

が同じグループへ分類

## 外部変数を導入した共起ネットワーク

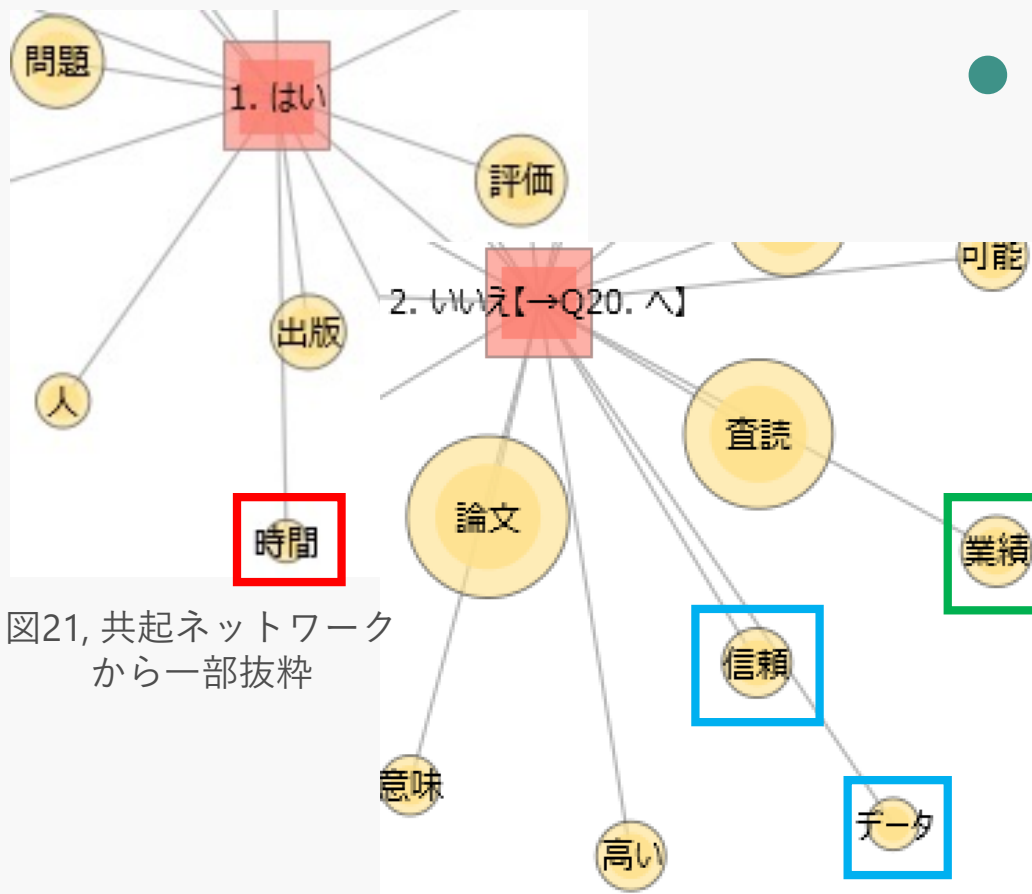


図21, 共起ネットワークから一部抜粋

図22, 共起ネットワークから一部抜粋

### ● 共起ネットワーク

- プレプリントの公開経験を外部変数として導入

### ● 公開経験あり

- 「時間」が分類  
回答例：“査読誌は出すのに時間がかかるが、プレプリントは速報性がある”

### ● 公開経験なし

- ! 「信頼」「データ」が分類  
回答例：“データの信頼性の担保をどうするのかという点”

- △ 「業績」が分類

## テキストマイニングのまとめ

### 1. 速報性への言及

- 「速報」「メリット」が同じクラスターへ分類  
プレプリントを公開した経験のある研究者が言及

### 2. プレプリントの信頼性への言及

- ! 「質」「信頼」「担保」などが同じクラスターへ分類  
プレプリントを公開した経験のない研究者が言及

### 3. プレプリントの内容判断についての言及

- △ 「読む」「人」「判断」が同じクラスターへ分類

### 4. プレプリントを業績として認めるかについての言及

- △ 「評価」「業績」「認める」が同じクラスターへ分類

## RQ1, 研究者のプレプリントに対する捉え方

### ● 利点

#### ● 速報性

- 速報性を理由にプレプリントを公開する研究者が多い
- 自由回答において「速報性」を利点として挙げる研究者が多い
- プレプリントの公開経験のある研究者が自由回答で多く言及

### ● 課題

#### ● 信頼性の保証

- 自由回答において信頼性の保証に言及した研究者が多い
- プレプリントの公開経験がない研究者が自由回答で多く言及

#### ● 内容の判断に注意が必要

- 信頼性の判断基準に本文や手法を挙げる研究者が多い
- 信頼性にDL数やオルトメトリクスなどの数値を挙げる研究者が少ない
- 内容の判断を読む人が行う必要があるという自由回答が多い

## RQ2, 研究者の考えるプレプリントの展望

- 分野ごとにプレプリントの発展状況は異なる可能性
  - 分野ごとに利用公開状況や今後の進展についての考え方が異なる
    - 数学:利用多・公開多・既に認められている多  
→既に十分認められているため,今後一定以上進展するとは言い難い
    - 心理学:利用多・公開多・今後認められる多  
→既に利用・公開ともに多いが,今後も一定以上進展すると予想される
    - 工学:利用少・公開少・今後認められる少  
→利用・公開ともに少なく,今後一定以上進展するとも言い難い
  - 全ての分野に「今後利用が進む」と回答した研究者が存在する  
→程度の差はあるが,全ての分野で今後も進む可能性がある

## RQ2, 研究者の考えるプレプリントの展望

- プレプリントを業績として扱うには注意が必要
  - 「プレプリントを業績として認めるべき」と回答した割合は3割強
    - 自由回答では「認めるべきでない」と判断される回答が22件
  - 「査読論文とは異なる方式で評価を行う必要がある」との回答も存在
    - 内容の審査・その後雑誌へ掲載されたか否か, などに基づく評価の可能性

## 結論

- プレプリントの利用は今後も進展する可能性がある
  - 速報性・先取性などの利点に注目する研究者が多い
  - 全ての分野に「今後利用が進む」と回答した研究者が存在する
- 業績として扱うか否かは議論が必要
  - 査読が存在しないことによる信頼性への不安感
  - 査読論文などとは異なる基準による評価の可能性
    - 内容の審査・引用数などの基準
    - その後雑誌へ掲載されたか否か
- プレプリントを利用する上でのリテラシーの周知が必要
  - 研究者以外もプレプリントに触れることが増えた
  - 査読を経っていないため、内容の判断は利用者が行う必要がある



## 参考文献

1. 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」, NISTEP RESEARCH MATERIAL, No. 301, 文部科学省科学技術・学術政策研究所. DOI: <https://doi.org/10.15108/rm301>
2. AMED. “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>, (参照 2022-10-22).
3. Polka, Jessica K; Penfold, Naomi C. "Biomedical preprints per month, by source and as a fraction of total literature". Zenodo. 2020-07-21. <https://zenodo.org/record/3955154#.YIK960zPIhF>, (accessed 2022-10-22).

## 参考文献

4. How a torrent of COVID science changed research publishing : in seven charts. nature. 2020, p.1-1.  
<https://www.nature.com/articles/d41586-020-03564-y>,  
(accessed 2022-10-22).
5. 吉田光男. COVID-19流行下でのプレプリントツイートに関する基礎調査. The 35th Annual Conference of the Japanese Society for Artificial Intelligence. 2021, p.1-4.  
[https://www.jstage.jst.go.jp/article/pjsai/JSAI2021/0/JSAI2021\\_1I4GS4c02/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/pjsai/JSAI2021/0/JSAI2021_1I4GS4c02/_pdf/-char/ja), (参照 2022-10-22).
6. なんでバズったの?: アスガルドアーキア論文掲載までのサイドストーリー. 日本微生物生態学会誌. 2021, 36(1), p.37-47.  
[https://www.jstage.jst.go.jp/article/jsmeja/36/1/36\\_37/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/jsmeja/36/1/36_37/_pdf/-char/ja), (参照 2022-10-23).

## 参考文献

7. 渡辺豪. “新型コロナ“ファクターX”は幻想? インド型変異株の「免疫を逃れる能力」と第5波のリスク”. AERAdot. 2021-06-09.  
<https://dot.asahi.com/aera/2021060800010.html?page=1>, (参照 2022-10-22).
8. 参考文献1に同じ
9. Nicholas Fraser et al. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. PLOS BIOLOGY. 2021, p.1-28.  
<https://journals.plos.org/plosbiology/article%3Fid%3D10.1371/journal.pbio.3000959?revisionLink=/plosbiology/article%3Fid%3D10.1371/journal.pbio.3000959%26rev%3D1>, (accessed 2022-10-22).
10. 池内有為, 大森悠生, 林和弘. “研究者によるオープンアクセスとプレプリント公開の実践状況と認識”. 第70回日本図書館情報学会研究大会. 宮城, 2022年10月29日. <https://doi.org/10.6084/m9.figshare.21432312>
11. 樋口耕一 2020 『社会調査のための計量テキスト分析 —内容分析の継承と発展を目指して— 第2版』 ナカニシヤ出版