

研究論文

Unpaywall を利用した日本におけるオープンアクセス状況の調査

The State of Open Access in Japan: An Analysis Using Unpaywall

西岡 千文^{1*} 佐藤 翔²

Chifumi NISHIOKA¹, Sho SATO²

¹ 京都大学附属図書館

Kyoto University Library

〒606-8501 京都市左京区吉田本町

E-mail: nishioka.chifumi.2c@kyoto-u.ac.jp

² 同志社大学免許資格課程センター

Center for License and Qualification, Doshisha University

〒302-8580 京都府京都市上京区新町通今出川上ル 同志社大学溪水館 315

E-mail: min2fly@slis.doshisha.ac.jp

* 連絡先著者 Corresponding Author

本稿は Unpaywall を利用して日本と世界のオープンアクセス（OA）状況を調査した。日本の調査では Unpaywall と Scopus に収録されている日本の著者による雑誌論文約 200 万件，世界の調査では Unpaywall に収録されている雑誌論文約 8,000 万件を対象とした。結果，日本と世界の OA の割合はそれぞれ 41.83%，29.77% であることがわかった。日本の割合の高さの要因は，過去に出版された論文の多くがブロンズであることである。近年の傾向としてゴールドの割合の上昇が観察された。機関リポジトリで公開されている論文の割合は 2000 年から 2010 年にかけて増加しているものの，2010 年以降は横ばいで約 5% である。2010 年以降の機関リポジトリのみで OA である論文の割合は約 1.5% である。グリーン OA を効率よく推進するためにも，各機関がそれぞれの OA の状況を継続的に把握できるような仕組みが必要である。

This article investigates the state of Open Access (OA) in Japan and the world using a dataset provided by Unpaywall. The analysis covers two million articles for Japan and 80 million articles for the world. We find that the percentages of OA articles in Japan and the world are 41.83% and 29.77%, respectively. A reason for the high percentage in Japan is that many articles published a long time ago belong to the bronze category. As a trend in recent years, the analysis reveals the percentage of articles in the gold category has increased. Regarding the contribution of institutional repositories (IRs) to OA, we observe the percentage of articles available in IRs increased from 2000 to 2010. However, the percentage has been flat, at 5%, since 2010. The percentage of articles that are OA only in IRs is 1.5%. To promote green OA efficiently, it is necessary to enable each institution to grasp its state of OA by e.g., developing a system that continuously monitors levels for each institution.

キーワード: オープンアクセス, 学術情報流通, オープンアクセスリポジトリ, 機関リポジトリ, 学術雑誌
open access, scholarly communication, open access repository, institutional repository, academic journal

1 はじめに

1990 年代以降，学術雑誌価格高騰への対応，研究成果の迅速かつ自由な共有の実現といった様々な動

機から，オープンアクセス（Open Access, OA）を推進する取り組みがなされてきた^[1]。2001 年に開催された Budapest Open Access Initiative (BOAI) は，OA の定義を，査読付きの学術雑誌論文について

て「インターネットにアクセスできることそれ自体を除く経済的、法的、技術的な障壁なく論文を利用できるようにすること」としている^[2, 1]。BOAIではOAを実現する手法として、セルフ・アーカイビングとOAジャーナルを挙げている^[2]。前者は、著者がOpen Archives Initiative (OAI) が定める標準(OAI-PMH等)に準拠したオープンな電子アーカイブ(リポジトリ)で論文の電子コピーを公開することを指し、グリーンOAと呼称される。後者は、著者がOAジャーナルに論文を掲載することでOAを実現するもので、ゴールドOAとされている。

2000年代以降のOAの進展を牽引してきたものとして、各国、各機関、各助成団体で制定されてきたOA方針が挙げられる^[3]。OA方針は各国、各機関に所属する、あるいは各助成団体から助成を受けている著者に、研究成果である論文をOAで公開することを義務付ける。世界に先駆けて、2005年に米国国立衛生研究所は、助成を受けた論文に対するパブリック・アクセス方針を公開した^[4, 5]。2010年代には、OAを推進する方策は広がりを見せ、OAは実現を疑われる理想から、実現されるであろう具体的な政策となっている^[3]。日本では、2009年に国立大学図書館協会による「オープンアクセスに関する声明」^[6]以降、OAの推進に関する議論が進んでいる。2011年に公表された第4期科学技術基本計画^[7]では、研究情報基盤の整備として「OAの推進」や「機関リポジトリの構築」について言及されている。2015年には内閣府の国際的動向を踏まえたオープンサイエンスに関する検討会から報告書^[8]が公表され、以降OAに関する政府提言が数多く発表されている。

機関単位でのOAの推進も行われてきた^[9]。2000年代にはハーバード大学といった欧米の機関でOA方針の策定が広まった。日本では、2015年に京都大学が機関リポジトリを通じたOAを義務化するOA方針^[10]を策定・公表した。2019年12月19日現在では、国内32機関でOA方針が定められている。

このように、これまでに政府や各機関の様々な方策によってOAが推進されてきた。今後さらにOAを推進するための戦略を策定するには、OAの現況を把握することが望ましい。これまでに世界の論文を対象としたOA状況の調査は行われている^[11, 12, 13]ものの、日本の論文を対象としたものは多くはなされていない。日本では機関ごとにOA方針の策定・OAの推進が行われてきたことから、機関

ごとのOA状況の調査が主に実施されてきたと考えられる。OAを推進するための国全体としての戦略の検討、各機関が国内での立ち位置を認識することを可能とするためにも、本稿では日本全体におけるOA状況の調査を行う。特に、日本のOA状況を、世界のそれと比較して概観する。調査では、Piwowarらの世界の論文を対象としたOA状況の調査^[11, 13]と同様に、データセットとしてUnpaywallのデータベースのスナップショットを利用した。UnpaywallはCrossref DOIが付与されている各論文の合法的にOAとなっている版を提供するウェブブラウザのプラグインである。多くの研究者に使用されており、2020年3月5日現在のユーザ数は204,896名である^[14]。また、2019年7月にはプラグインを通じて、約300万件の論文のOA版のリクエストがあったことが報告されている^[13]。よって、OAとなっている論文の発見可能性という観点からも、Unpaywallを使用した調査には意義があると考えられる。ゴールド、ハイブリッド、ブロンズ、グリーン、クローズドといったUnpaywallで与えられているOA種別の観点からの調査とともに、機関リポジトリ等のリポジトリによるOAという観点からも調査を行った。

本稿の構成を以下に示す。次章では、関連研究として様々なOA種別とこれまでに実施されたOA状況の調査を挙げる。3章では調査の手法として、本稿で使用するデータセットと各論文のOA種別を特定する方法を示す。4章では調査の結果を示し、続く5章では結果を考察する。6章を本稿の結びとする。

2 関連研究

本章では、OA状況の調査についての関連研究を示す。2.1節では、これまでのOA状況の調査で使われてきた様々なOA種別について述べる。続く2.2節では、これまでに実施された日本と世界のOA状況の調査を挙げる。

2.1 様々なOA種別

BOAIで示されているグリーンOA、ゴールドOAをはじめとして、OAの実現には様々な手段が存在する。このことから、これまでのOA状況の調査では、調査ごとに様々なOA種別が使用されてきた。Piwowarらの調査^[11]での文献レビューをもとにして、様々なOA種別を下記に示す。

- **ゴールド**. OA ジャーナルで出版されている論文を指す.
- **ハイブリッド**. 購読型雑誌で出版されているが、論文掲載料 (Article Processing Charge, APC) の支払いで OA となった論文を指す.
- **グリーン**. 購読型雑誌で出版されており出版者では OA となっていないが、機関リポジトリ、プレプリントサーバ等のリポジトリで OA となっている論文を指す [15, 16].
- **Libre**. 読むことができる権利に加えて、クローリング、インデキシング、アーカイブ等のために再利用できる権利が与えられている論文を指す [17]. BOAI^[2] での OA の定義に近い.
- **Gratis**. Libre では再利用できる権利が与えられているのに対して、読むことができる権利のみが与えられている論文を指す [17].
- **遅延 OA**. 購読型雑誌に掲載されており、一定期間後に出版者で OA となる論文を指す. Laakso らの調査 [18] によると一定期間後に出版者で OA となる論文のうち 77.8% が出版から一年以内に OA となっている.
- **ASN**. ResearchGate^[19] や Academia^[20] といった商業のアカデミック・ソーシャル・ネットワーク (ASN) で公開されている論文を指す.
- **ブラック** [21]. Sci-Hub といった学術論文の海賊版サイトに公開されている論文を指す. OA 種別としては存在するものの、違法であることから多くの OA 状況の調査で対象とはなっていない. これらは Rouge OA や Robin Hood OA^[22, 23] とも呼称される.
- **ブロンズ**. ゴールド、ハイブリッドに該当しないが、出版者で OA となっている論文を指す.
- **プラチナ/ダイヤモンド** [24, 25]. APC が無料である OA ジャーナルで出版されている論文を指す. 上記のゴールドに含まれている. 出版に要する費用は多くの場合、学協会によって負担されている.

本稿の調査では、Piwowar ら [11] に倣って、ゴールド、ハイブリッド、ブロンズ、グリーンの OA 種

別を利用する (3.2 節参照). この OA 種別は、国外での OA 状況の調査 [26] 等、近年最もよく利用されている. 遅延 OA については、Piwowar らの最新の調査 [13] でも扱われているが、本稿では扱わない. ASN については、一部の OA の調査 [22, 27] では一つの OA の手法として捉えられている. しかし、Jamali^[28] によると ResearchGate で公開されている購読型雑誌の論文のうち約半数が出版者の規定に反していることが判明している. また、永続的な公開といった観点から、ASN での OA には懸念が抱かれる [29, 30]. これらの理由から、Piwowar らの調査 [11, 13] と同様に、本稿では ASN での OA は扱わない. また、ブラックについては、その手法に法的な懸念があることから、本稿では扱わない. 近年 APC の値上げが進んでいる [31] ことから、プラチナ/ダイヤモンドは注目すべき OA 種別であるが、別稿に譲る.

本稿ではさらに、DOAJ への掲載とライセンスの付与という観点からゴールドを細分化した分析を実施する. DOAJ (Directory of Open Access Journals) [32] は、OA ジャーナルを「全コンテンツに無償でアクセス可能である」かつ「査読・編集によって高品質基準を満たしている」学術雑誌と定義しており、これらの要件を満たす学術雑誌のリストを公開している. 捕食ジャーナル、ハゲタカジャーナル [33] といった問題が近年注目を集める [34] 中、DOAJ は OA ジャーナルのホワイトリストとして使用されている. また、BOAI による OA の定義 (1 章参照) [2, 1] や Suber による Libre の定義 [17] にもあるように、コンテンツの再利用という観点からライセンスを付与することは重要である. これらのことから、DOAJ への掲載とライセンスの付与という観点からの分析を実施する.

また本稿では、リポジトリによる OA に注目した調査を実施する. Harnad ら [15, 16] や Piwowar ら [11] による調査でのグリーンは、購読型雑誌に掲載されておりリポジトリで OA となっている論文を指す. よって、「論文が出版者で OA である」かつ「リポジトリでも公開されている」際には、グリーンであるとは認識されない. さらに、これらのグリーンの定義では機関リポジトリとその他リポジトリを区別していない. 機関リポジトリの「世界に対する大学の貢献を形作る新しいチャネル」 [35, 36] という役割に注目した際には、OA ジャーナルに掲載されている論文が機関リポジトリで OA となっていること

は意義がある。このことから、機関リポジトリにおける OA 状況は大きな関心である。

2.2 OA 状況の調査

OA への関心の高まりから、数々の OA 状況の調査が実施されてきた。本節では、日本と世界における OA 状況の調査についてそれぞれ述べる。

2.2.1 日本の OA 状況の調査

日本の論文全件を対象とした大規模な OA 状況の調査はなされていないが、関連する調査を下記に挙げる。

福澤^[37]は、Scopus を利用して、日本、米国、ドイツ、フランス、英国、中国、韓国といった主要国の論文発表の特徴をジャーナルに注目して分析している。研究では、「OA ジャーナル/それ以外のジャーナル」という観点からも分析しており、2010～2012 年の期間で日本の機関の著者により発表された論文のうち、自国の OA ジャーナルから出版されたものが 4.2%、他国の OA ジャーナルから出版されたものが 7.4%を占めることが明らかになっている。

池内ら^[38]は日本の研究者を対象として、研究データの公開を中心としたオープンサイエンスの実態についてのアンケート調査を実施している。回答者 1,398 名のうち 70.9%が OA の論文があると回答している。

尾鷲ら^[27]は、国立環境研究所における OA について、「研究成果のうち OA になっているものはどの程度か」を探る OA 度調査と「OA のために必要となる APC の価格」を探る OA 化調査を実施している。調査では、所内で運用されている研究成果データベースから抽出した 2011～2016 年度の論文を対象としている。対象となる各論文について、OA であるか否か、OA である際にはさらにゴールドであるかグリーンであるかというようにその手段にまで言及している。グリーンである場合には、リポジトリでの OA であるか、ASN での OA であるか、といったプラットフォームの調査も実施している。ASN での OA もグリーンとして扱う点は、本稿とは異なる。結果として、2016 年度では 7 割近くの論文が何かしらの手段で OA となっていることが判明した。ゴールドの割合が大きく上昇していることが観察された。

国立大学図書館協会オープンアクセス委員会は、2016～2017 年にかけて加盟館に対してオープンアクセスへの取り組みに関する実態調査^[39]を実施した。その中で、コンテンツ種別（e.g., 学術雑誌論文、研究費研究成果報告書）ごとに機関リポジトリへの登録状況を調査している。結果は大学の規模（学部数で算出）ごとに示されている。2014～2015 年の学術雑誌論文（海外）のリポジトリでの捕捉率については、8 学部以上ある大学では 10%程度、5～7 学部の大学では 2～3%程度、2～4 学部の大学では 7～10%程度、単科大学では 7%程度であった。

大学図書館コンソーシアム連合（JUSTICE）は、日本の著者の論文件数、OA ジャーナルで出版された論文の割合および APC 支払推定額の調査を実施しており、その結果は小陳ら^[40]によって報告されている。調査では Web of Science のデータを使用しており、2012～2014 年に出版された論文を対象としている。期間中の論文件数は年間約 65,000 件とほぼ横ばいである。OA ジャーナル掲載論文件数は年々増加しており、全論文件数に対する割合は、2012 年が 11.6%、2013 年が 13.4%、2014 年は 14.9%になっている。

Archambault ら^[22]の調査（後述）では、国別の OA 状況の分析を実施しており、日本についても触れられている。2008～2013 年に出版された Scopus に収録されている論文を対象としており、日本の論文の OA の割合は 44.2%であり、EU28/ERA の加盟国平均の 51.3%や世界平均 47.1%と比較すると低い。なお、これらの数値は、論文の OA 版を収集するハーベスト用のエンジンの見落としを考慮して調整が行われている。

2.2.2 世界の OA 状況の調査

世界の論文を対象とした OA 状況の調査は数多く実施されている。

Martín-Martín ら^[12]は、2009～2014 年に出版された Web of Science に収録されている論文を対象とした OA 状況の調査を Google Scholar を使用して実施している。結果、調査対象の 55%の論文の OA 版が Google Scholar で見つけられることが判明した。23%が出版者で、17%がリポジトリで、40%が ResearchGate といったその他のサイトで OA であることがわかった。分野ごとの調査も実施しており、OA の割合が高い分野として自然科学、医学、低い

分野として社会科学、工学、人文社会科学を挙げている。地域ごとの調査では、欧州、アメリカ、ブラジルで高い OA の割合（36～50%）であるのに対して、日本を含めたアジア諸国では 18～36%であることを明らかにした。

Archambault らは^[22]、Scopus に収録されている 1996～2013 年に出版された論文を対象として、OA 状況の調査を実施している。この調査では、論文の OA 種別としてゴールド、グリーン、その他 OA（ブロンズや ASN での OA 等を含む）が採用されている。結果として、以前と比較してグリーンの伸び率が小さくなってきていることとゴールドが増加していることが報告されている。2008～2013 年に出版された論文を対象として、分野別、国別の調査も実施している。OA の割合が高い国として、オランダ、クロアチア、エストニア、ポルトガル、スイス、ブラジルが挙げられている。OA の割合が高い分野としては、一般科学・技術、生物医学、数学・統計学、生物学が挙げられている。

最近の大規模な調査としては、Piwowar らの調査^[11]が挙げられる。この調査は、尾城らの論考^[41]でも取り上げられている。調査では、Crossref DOI をもつ全ての論文 66,560,153 件のうち 100,000 件、Web of Science に収録されている Crossref DOI をもつ論文 8,083,613 件のうち 100,000 件、2017 年のとある週に Unpaywall のユーザがアクセスした論文 213,323 件のうち 100,000 件をサンプルとして使用している。また、詳細な OA 状況の把握のために、各論文をゴールド、ハイブリッド、ブロンズ、グリーン、クローズドのいずれかに分類している（2.1 節参照）。Crossref DOI をもつ論文のサンプルについては、世界で公開されている論文全件での OA 状況进行分析のために使用している。結果、27.9%の論文が OA となっていることが判明した。内訳としては、ゴールドが 3.2%、ハイブリッドが 3.6%、ブロンズが 16.2%、グリーンが 4.8%である。さらに、論文の出版年別、出版者別、分野別の分析も実施している。出版年別の調査では、出版年が最近であるほど、OA の割合が高いことが判明している。Web of Science に収録されている Crossref DOI をもつ論文のサンプルは、論文の OA が被引用数に与える影響を調査するために利用している。結果、グリーン、ハイブリッド、ブロンズであるものについては、平均よりも多く引用されていることが判明した。対してゴールドである論文は、平均やクローズドである

論文よりも被引用数が小さい。Unpaywall のユーザがアクセスした論文のサンプルは、ユーザが利用する論文における各 OA 種別の割合を計算するために利用されている。結果として、ユーザがアクセスした論文のうち 47.0%が OA であることが判明した。

Piwowar らは 2019 年にも OA 状況の調査^[13]を実施している。この調査では、先行研究^[11]と同様に、ゴールド、グリーンといった各 OA 種別の割合や、Unpaywall のユーザにアクセスされている論文での OA の割合を分析している。さらに発展させて、グリーンとブロンズのタイムラグ（OA になるまでに要する時間）についても分析を実施している。グリーンについては、プレプリントの登録はごく僅かであり、多くの場合に出版から約 1 年後にリポジトリで公開されていることが判明した。回帰分析によって将来の予測も実施しており、2025 年の OA の割合は 44%であるとしている。

3 手法

本章は OA 状況の調査手法について述べる。3.1 節では利用したデータセットについて、3.2 節では各論文の OA 状況の判断方法について詳述する。

3.1 データセット

Piwowar らによる調査^[11, 13]と同様、本稿でも Unpaywall を利用して調査を実施する。Unpaywall が利用するデータベースは公開されており、本稿では 2019 年 11 月 22 日版を利用した。スナップショットには、Crossref DOI が付与されている 109,905,121 件のレコードのメタデータならびに OA に関する情報が収録されている。分析では日本と世界の OA 状況を比較するが、それぞれ下記を調査対象とする。

- 日本。Scopus で日本に位置する機関 2,611 件とそれらの機関の Affiliation ID を取得する。Scopus の Affiliation Retrieval API^[42]を利用して、それぞれの機関の構成員が著者となっている DOI が付与された論文を計 3,054,454 件（異なり数）抽出した。それらのうち、Unpaywall のスナップショットに収録されているかつ種別が `journal-article`（雑誌論文）である論文 2,000,897 件を調査対象とする。

- 世界. Unpaywall のスナップショットに収録されていて、種別が `journal-article` (雑誌論文) である論文 79,534,697 件を調査対象とする。これには上記の日本の論文も含まれる。

本稿では分野ごとの OA 状況の調査 (4.2 節参照) も実施する。しかし Unpaywall は論文の分野に関するメタデータを収録していない。よって、本稿では ESI Master Journal List^[43] (2019 年 8 月版) を使用することで、各論文の分野を特定した。ESI Master Journal List は、Essential Science Indicators に収録されている学術雑誌のリストである。リストには 12,001 件の各学術雑誌のタイトル、ISSN、EISSN、分野が収録されている。分野は 22 件存在し、各学術雑誌につき 1 件の分野が与えられている。分野は論文単位ではなく学術雑誌単位に与えられているものであるが、分野ごとの OA 状況を大まかに把握できると考える。ISSN、EISSN によって紐付けを行い、論文の分野を特定した。1 件の ISSN に複数の学術雑誌が紐付けられていることなどを理由として、論文に複数の分野が与えられるケースが発生した。その際には、1 を分野数で割った値を各分野の重みとする。

3.2 論文の OA 種別の特定

本稿では、下記の 2 つの観点から OA 状況を調査する。

- OA の全体像 (Unpaywall の OA 種別を利用)
- リポジトリによる OA

本節では、上記の 2 つの観点からの OA 種別とその特定方法について詳述する。

3.2.1 OA の全体像

Unpaywall の各論文には、フィールド `oa_status` が存在し、`gold`, `hybrid`, `bronze`, `green`, `closed` いずれかの値が与えられている。この OA 種別は、Piwowar らの調査^[11] で初めて使用され、現在までに欧州諸国での OA のモニタリング等、様々な調査で利用されている^[44, 26]。これらの OA 種別の付与方法は公開されており^[45]、表 1 にまとめる。なお、フィールド `oa_status` が存在しないあるいはフィールドが空値である論文が含まれていることが判明し

たため、本稿では新たに不明という OA 種別を追加している。

表 1 にあるとおり、論文が掲載されている学術雑誌の種別 (OA ジャーナルまたは購読型雑誌) が、OA 種別の決定に大きな役割を果たしている。OA ジャーナルの判定方法についても公開されており^[46]、判定方法を図式化したものを図 1 に示す。

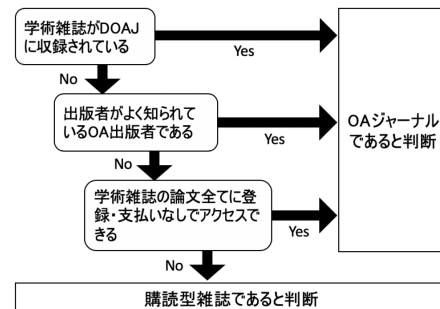


図 1: Unpaywall における OA ジャーナルの判定方法。

また、Unpaywall では各論文がリポジトリで公開されているか調査することで、グリーンの判定を行っている。リポジトリについてであるが、学術機関リポジトリデータベース (IRDB)^[47] に登録されている機関リポジトリ 631 件のうち、Unpaywall に含まれているリポジトリは 76 件である。この数値は、2019 年 7 月 10 日時点での IRDB に登録されている機関リポジトリの URL のネットワークロケーションのリストと Unpaywall が日本の論文を公開していると判断しているリポジトリの URL (3.2.2 項参照) のネットワークロケーションのリストを照合することによって算出した。多くの機関リポジトリは主として紀要や学位論文など灰色文献 (グレイリテラチャー) の流通に貢献している^[48] ことから、76 件はそこまで小さい値ではないと考える。なお、IRDB には登録されていないが、Unpaywall が日本の論文を公開していると判断するリポジトリ (海外のリポジトリ、誤ってリポジトリとして捉えられているものも含む) は 2,918 件存在する。

いずれにせよ全ての日本のリポジトリが対象となっているわけではなく、結果で示す日本の論文のグリーンの割合は実際よりも低いといえるだろう。本稿ではリポジトリによる OA に注目した調査も実施するが (3.2.2 項参照)、それらの結果の数値も低くなるといえる。

また、本稿ではゴールドの詳細な OA 種別を適宜

表 1: 「OA の全体像」の調査における OA 種別

OA 種別	OA 種別が付与される条件
ゴールド (gold)	OA ジャーナルに掲載されている論文
ハイブリッド (hybrid)	購読型雑誌に掲載されているが、CC-BY 等のオープンライセンスが付与されている論文
ブロンズ (bronze)	上記いずれにも該当しないが、出版者で OA となっている論文
グリーン (green)	上記いずれにも該当せず、リポジトリのみで OA となっている論文
クローズド (closed)	OA でない論文
不明	フィールド <code>oa_status</code> が存在しない、または空値である論文

利用する。図 1 にあるように、「DOAJ への収録」、「出版者がよく知られている OA 出版者」、「学術雑誌に掲載されている論文全件に無料でアクセス可能」のいずれかに合致すれば、学術雑誌は OA ジャーナルであると判断されている。2.1 節でも述べたとおり、近年捕食ジャーナル、ハゲタカジャーナルといったといった問題が注目される中、DOAJ への収録は重要になっている。また BOAI で述べられている「できる限り制限のない利用条件」を実現するためには、ライセンスを明記することが重要である。これらのことから、ゴールドを下記のとおり 4 つの OA 種別に分類して詳細な調査を実施する。なお、本稿では「DOAJ 掲載論文」は「DOAJ に収録されている OA ジャーナルに掲載されている論文」を指す。

- DOAJ 掲載論文・ライセンス付
- DOAJ 掲載論文・ライセンスなし
- その他 OA ジャーナル掲載論文・ライセンス付
- その他 OA ジャーナル掲載論文・ライセンスなし

3.2.2 リポジトリによる OA

表 1 にあるように、Unpaywall の OA 種別でのグリーンは「リポジトリのみで OA となっている論文」を指す。よって、出版者で OA となっている論文は、リポジトリで公開されていたとしても考慮されない。しかし機関リポジトリの「世界に対する大学の貢献を形作る新しいチャンネル」^[35, 36]や「機関の研究成果やその他の学術および教育リソースを発信」^[49]という役割に注目すると、出版者で OA となっている論文も機関リポジトリに収録されることが望ましい。さらに、著者の OA に対する意識を把握するためにも、プレプリントサーバ等その他のリ

ポジトリでの OA についても調査することが求められる。これらのことから本稿では、リポジトリによる OA に注目した調査を実施する。

表 2 にこの調査で利用する OA 種別とその付与方法を記す。Unpaywall の各論文には `oa_locations` というフィールドが存在しており、論文が OA である際には 1 件以上のオブジェクトが格納されている。各オブジェクトには、OA となっているサイトについての情報が格納されている。具体的には `host_type`, `url` といったフィールドが存在しており、「出版者での OA またはリポジトリでの OA であるか」といった情報や「その論文の OA 版の URL」が格納されている^[50]。これらの情報を利用して、OA 種別を付与する。

なお、表 2 の 2 行目以下の OA 種別については、日本の論文のみを調査対象とする。また、この OA 種別における「機関リポジトリ」は「日本の機関リポジトリ」を指し、海外の機関リポジトリはその他リポジトリとして認識される。表 2 の 2・3 行目にある URL のネットワークロケーションは、Python のモジュール `urllib.parse`^[51] を使用して特定する。調査では、ネットワークロケーションが `.jp` で終わるリポジトリを機関リポジトリとする。日本の論文を多く公開しているネットワークロケーションが `.jp` で終わるリポジトリ上位 30 件を調査したところ、30 件中 29 件が機関リポジトリであった。ネットワークロケーションが `.jp` で終わるリポジトリは 448 件見つかり計 53,441 件の日本の論文を公開している。そのうち上位 30 件のリポジトリで公開されている論文は計 49,392 件であり、全体の 92.42% を占めている。よって、この調査で報告される機関リポジトリによる OA の割合は、実際よりも低いことはあっても、高いことはないといえるだろう。

表 2: 「リポジトリによる OA」の調査における OA 種別.

OA 種別	OA 種別が付与される条件
リポジトリで OA である論文	フィールド <code>oa_locations</code> に, <code>host_type</code> が <code>repository</code> であるオブジェクトが 1 件以上格納されている論文
機関リポジトリで OA である論文	フィールド <code>oa_locations</code> に, <code>host_type</code> が <code>repository</code> であるかつ <code>url</code> のネットワークロケーションが <code>.jp</code> で終わるオブジェクトが 1 件以上格納されている論文
機関リポジトリ以外のリポジトリで OA である論文	フィールド <code>oa_locations</code> に, <code>host_type</code> が <code>repository</code> であるかつ <code>url</code> のネットワークロケーションが <code>.jp</code> で終わらないオブジェクトが 1 件以上格納されている論文
機関リポジトリとその他リポジトリで OA である論文	上記の「機関リポジトリで OA である論文」・「機関リポジトリ以外のリポジトリで OA である論文」両方に該当する論文
機関リポジトリでのみ OA である論文 (出版者でも OA ではない論文)	上記の「機関リポジトリで OA である論文」に該当し, フィールド <code>oa_locations</code> に格納されているオブジェクトが 1 件のみである論文

4 結果

本章では, 日本における OA 状況の調査結果を示す. 4.1 節では, 日本と世界の論文における OA 状況について述べる. 続いて, 分野 (4.2 節), 出版年 (4.3 節) といった観点から, OA 状況を報告する. 4.4 節では論文の公開に利用されているリポジトリについて述べる.

4.1 節, 4.2 節, 4.3 節では 3.2 節で示した OA の全体像とリポジトリによる OA という 2 つの観点からの分析結果を別々に示す.

なお, 本章では結果の多くを図で表しているが, 図のもととなったデータについては付録に示す (本稿の電子版参照).

4.1 全体

4.1.1 OA の全体像

日本と世界の OA 状況を表 3 に示す. OA である論文の割合は, 日本と世界それぞれで 41.83%, 29.77% であることから, 日本のほうが高い. 理由としては, 日本のブロンズの割合の高さが挙げられる. その他の OA 種別の割合については, 日本と世界で大きな差は存在しない.

また, 3.2.1 項で示したとおり, DOAJ への掲載とライセンスの付与という観点から, ゴールドを 4 つの OA 種別に分類した結果も表 3 に示す. ここでも日本と世界で大きな差は観察されない. ゴールドの中では, 「DOAJ 掲載論文・ライセンス付」と「その他 OA ジャーナル掲載論文・ライセンスなし」に分類されるものが多数を占める. DOAJ の収録要件としてライセンスの明示が挙げられている [52]. こ

の要件が論文を DOAJ 掲載論文とその他 OA ジャーナル掲載論文に分割しているといえる.

なお本稿では, 日本の機関の構成員が少なくとも一名以上著者となっていて, かつ Scopus に収録されている論文を日本の論文としている. Scopus に収録されるには, 査読体制, 被引用数等の基準 [53] を満たした学術雑誌に掲載されなければならない. よってここでの日本の論文には, ハゲタカジャーナル掲載論文は含まれていない可能性が高い. 一方, 世界の論文については, Crossref DOI が付与されている論文全てが対象となっているので, ハゲタカジャーナル掲載論文が含まれているだろう.

表 3: 日本と世界の OA 状況.

OA 種別	日本 (%)	世界 (%)
OA	41.83	29.77
ゴールド	7.33	7.67
DOAJ 掲載論文・ライセンス付	5.02	5.22
DOAJ 掲載論文・ライセンスなし	0.34	0.66
その他 OA ジャーナル掲載論文・ライセンス付	0.29	0.71
その他 OA ジャーナル掲載論文・ライセンスなし	1.68	1.09
ハイブリッド	3.09	3.37
ブロンズ	24.37	10.45
グリーン	7.04	8.28
クローズド	58.16	70.08
不明	0.01	0.15

4.1.2 リポジトリによる OA

続いて, 3.2.2 項で示した手法を用いて, 日本と世界のリポジトリによる OA 状況を調査した. 結果

を表4の1行目に示す。表3では世界のグリーンの割合のほうが僅かに高かったが、リポジトリでOAである論文の割合は日本のほうが高い。このことから、日本のほうが出版者でOAとなっている論文をリポジトリによく登録しているといえる。リポジトリの種別などに注目した分析の結果を、表4の2行目以下に示す。圧倒的に多くの論文が、機関リポジトリ以外のリポジトリで公開されていることがわかる。なお表4の1~4行目は、出版者でOAとなっている論文も含む。表4の5行目に示す機関リポジトリのみでOAである論文は、0.83%とごく僅かである。

4.2 分野

異なる学術分野では出版についても様々な慣習をもっており、OA状況には大きな差異があることが想定される。このことから、分野ごとにOA状況を探る。各論文の分野は、3.1節で述べたとおり、ESI Master Journal List を使用して特定した。

4.2.1 OAの全体像

図2と図3に、日本と世界における分野ごとのOA状況を示す。図の右端にある数値は各分野の論文件数である。

ほとんどの分野において、日本の各OA種別の割合は、世界の割合よりも高い。物理学、生物学・生化学、神経科学・行動科学、微生物学、環境学・生態学の分野においては、世界の割合のほうが高い。神経科学・行動科学については、差が6.73%存在する。日本における当該分野でのブロンズとグリーンの割合の小ささが原因である。生物学・生化学では差が5.09%存在しており、日本における当該分野でのハイブリッドとグリーンの割合の小ささが原因となっている。物理学では差が4.70%存在しており、日本における当該分野でのグリーンの割合の小ささが理由である。その他分野では差はいずれも2%未満であった。

OAの割合が高い分野としては、学際的領域と宇宙科学が挙げられる。いずれも80%以上という高い割合である。なお、3.1節で述べたとおり、分野は学術雑誌単位で与えられるものである。よって学際的領域は「論文が学際的な研究を扱っている」というわけではなく、むしろ「論文が掲載されている学術

雑誌が様々な分野を対象としていて学際的である」というほうが近い。学際的領域についてはゴールド、宇宙科学についてはブロンズとグリーンによって、OAの割合が高くなっている。日本ほど高い数値ではないものの、世界における学際的領域と宇宙科学でも同様の傾向が観察される。学際的領域については、メガジャーナル^[54]の影響を受けている。日本においては学際的領域の論文が44,571件(図2参照)存在するが、そのうち12,613件がPLOS ONE(実際の集計では6,351件がPLoS ONE, 6,262件がPLOS ONEであったが、同一の学術雑誌であるので合算)、9,428件がScientific Reportsに掲載されているものである。宇宙科学については主要な学術雑誌であるThe Astrophysical Journal, Publications of the Astronomical Society of Japan, Monthly Notices of the Royal Astronomical Societyに掲載されている論文の多くがブロンズであることによって、ブロンズの割合が高くなっている。

続いて図2と図3からゴールドを分野ごとに抽出し、DOAJへの掲載とライセンスの付与という2つの観点からの詳細なゴールドのOA状況を調査した。結果を図4(日本)と図5(世界)に示す。ほとんどの分野で、「DOAJ掲載論文・ライセンス付」が高い割合を占めていることがわかる。例外としては、日本における薬理学・毒性学と農学が挙げられ、「その他OAジャーナル掲載論文・ライセンスなし」が半数以上を占めている。

4.2.2 リポジトリによるOA

リポジトリによるOAの全体像を探るために、各分野ごとにリポジトリでの公開状況を探る。プレプリントサーバであるarXivをはじめとして、SSRN, bioRxivと様々なリポジトリが分野ごとに発展してきた。このことから、分野ごとに異なる結果が観察されることが期待される。

図6に、日本と世界でのリポジトリで公開されている論文の割合を分野ごとに示す。日本においては、学際的領域と宇宙科学で高い割合が示されている。図2にあるとおり学際的領域ではメガジャーナルの影響からゴールドが多数を占めており、相対的にはグリーンの割合は高くない。このことは、学際的領域でゴールドである論文の多くが、リポジトリにも登録されていることを示す。世界においては、学際的領域、微生物学、宇宙科学がいずれも4割程

表 4: 日本と世界のリポジトリによる OA 状況.

OA 種別	日本 (%)	世界 (%)
リポジトリで OA である論文	19.90	17.57
機関リポジトリで OA である論文	2.36	—
機関リポジトリ以外のリポジトリで OA である論文	18.71	—
機関リポジトリとその他リポジトリで OA である論文	1.17	—
機関リポジトリでのみ OA である論文 (出版者でも OA ではない論文)	0.83	—

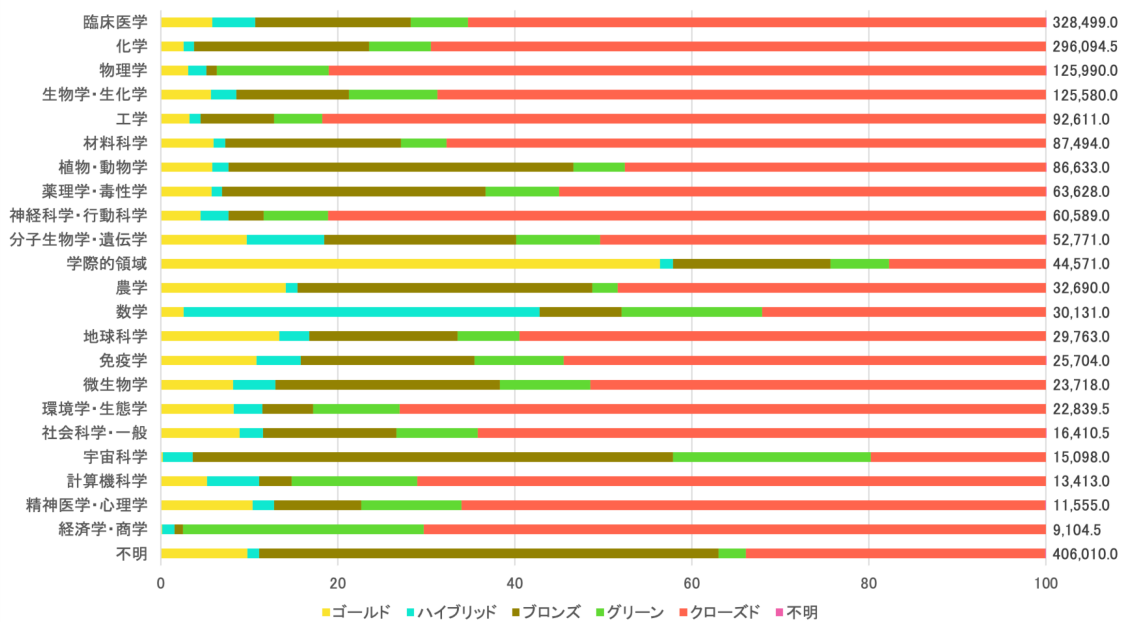


図 2: 日本における分野ごとの OA 状況.

度であり、他の分野よりも高い数値を示している。グリーンが進んでいない分野としては、日本と世界共通して、化学、工学、材料科学が挙げられ、いずれも 1 割程度である。

リポジトリの種別などに注目したより詳細な結果を、図 7 に示す。図 6 より学際的領域と宇宙科学の論文が多くリポジトリで公開されていることが判明したが、その多くは機関リポジトリ以外のリポジトリで OA となっていることがわかった。特に学際的領域はメジャーナルの影響を受けており、それらの多くがライセンスを明記していることから、Semantic Scholar や PubMed Central といった様々なサービスで公開されている。なお、Semantic Scholar や PubMed Central をリポジトリとして捉えることには議論があると思われる (4.4 節参照) が、Unpaywall ではリポジトリの一つとして捉えられているため、そのままの結果を示している。機関

リポジトリでの公開が進んでいる分野として、学際的領域と経済学・商学が挙げられる。学際的領域については、メジャーナルの論文を中心にライセンスが付与されたゴールドの論文が多いことから、機関リポジトリへの登録が進展していることが理由として考えられる。経済学・商学については、機関リポジトリで公開されている論文が特定の学術雑誌に偏っているということは観察されなかった。

機関リポジトリでの公開が進んでいない分野として、不明、農学、薬理学・毒性学が挙げられる。理由の一つとして、これらの分野ではゴールドでもライセンスが付与されていない論文が多く (図 4 参照)、機関リポジトリでの公開が進みにくいことが挙げられる。

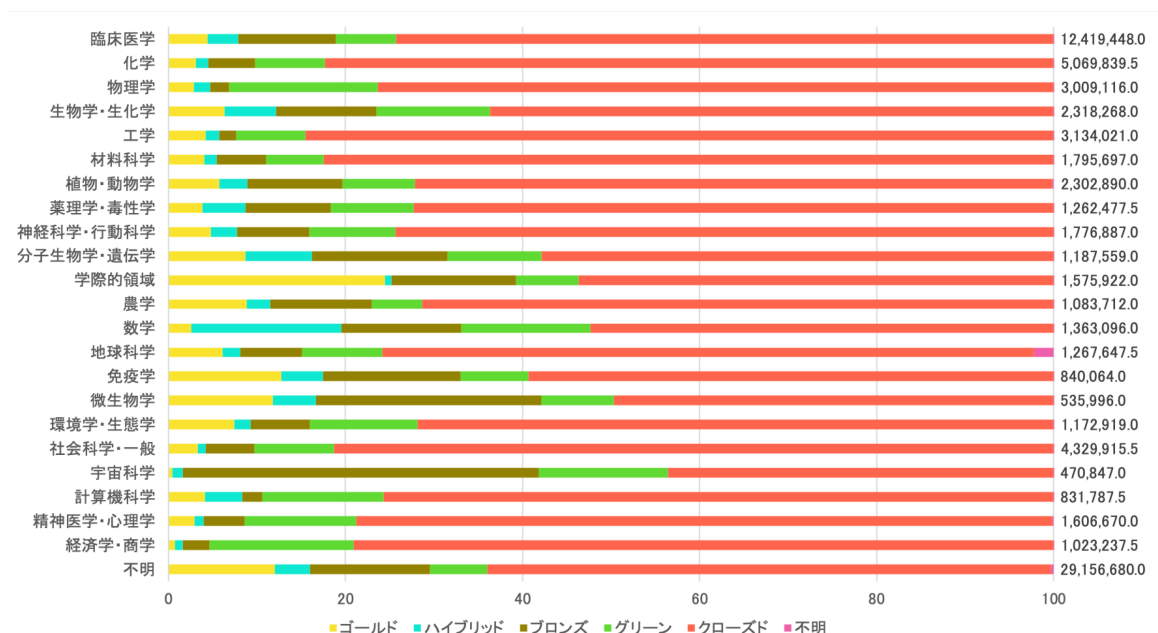


図 3: 世界における分野ごとの OA 状況.

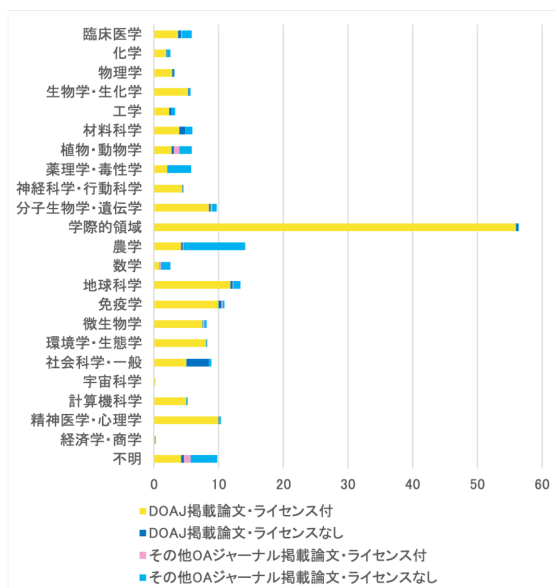


図 4: 日本における分野ごとのゴールド OA の状況.

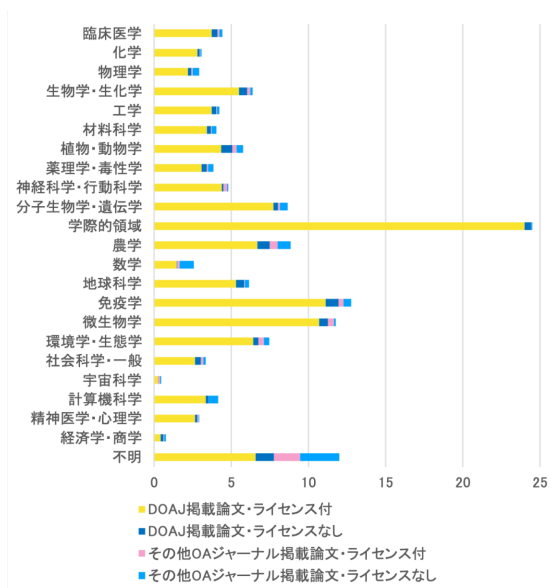


図 5: 世界における分野ごとのゴールド OA の状況.

4.3 出版年

過去 30 年間で学術情報流通は大きく変化した。近年では、OA の需要・認知が高まり、様々な OA ジャーナルが興隆している。また、出版者によっては論文の OA を一定期間禁止するエンバゴ期間を設けている。このことから、論文の OA 状況は

出版年によって大きく変化すると考えられる。本節では、論文の出版年別に OA 状況を調査する。各論文の出版年は Unpaywall に登録されているメタデータより取得した。対象とする出版年は 1990～2018 年であり、この期間に出版された日本の論文は 1,782,243 件（日本の論文全件の 89.07%）、世界の論文は 61,949,966 件（世界の論文全件の 77.89%）

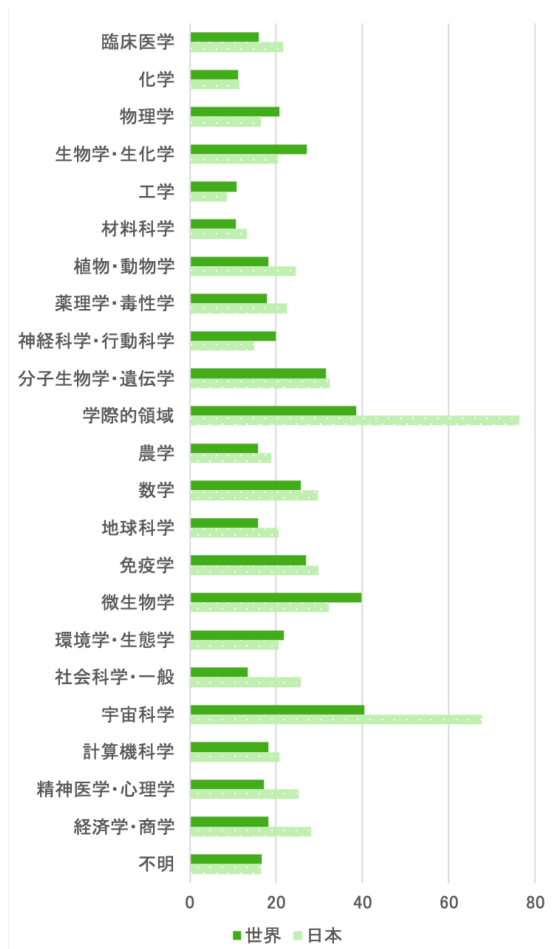


図 6: 分野ごとのリポジトリによる OA 状況.

である．以下では，OA の全体像とリポジトリによる OA という異なる 2 つの観点から結果を述べる．

4.3.1 OA の全体像

図 8 と図 9 に，日本と世界における出版年ごとの各 OA 種別の割合（数値は左軸に対応）と全論文件数（数値は右軸に対応）を示す．日本では，直近の 10 年間に出版された論文に注目すると，概ね 40% の論文が OA になっていることがわかる．Archambault ら^[22] は 2008～2013 年を対象とした調査で日本の OA の割合を 44.2% と報告しているが，その結果とも概ね一致する．なお，Archambault ら^[22] は ResearchGate^[19] 等の ASN (Academic Social Network) での OA もカウントしていることから，彼らが算出した OA の割合のほうが若干高い．Web of Science を利用して調査を行なった小陳ら

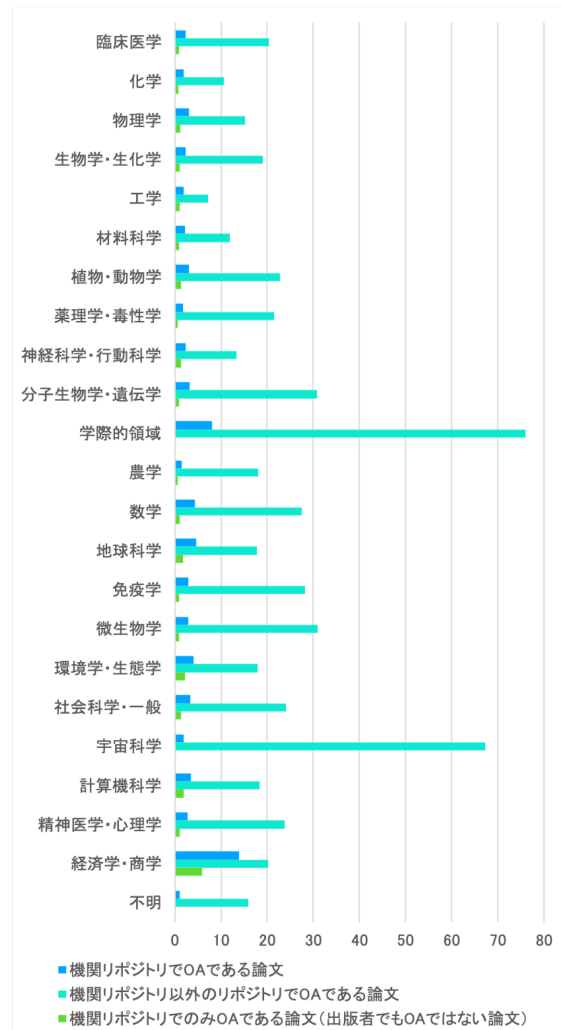


図 7: 日本における分野ごとのリポジトリによる OA 状況.

[40] は，2012～2014 年に出版された論文件数はいずれの年も 65,000 件程度であったと報告している．本稿で使用した Scopus は Web of Science よりも収録範囲が広いので，1 割程度論文件数が多い．内訳に注目すると，特にゴールドの割合が近年顕著に伸びている．小陳ら^[40] は，OA ジャーナル掲載論文件数 (i.e., ゴールド) の割合を 2012 年が 11.6%，2013 年が 13.4%，2014 年が 14.9% であると報告している．図 8 では 2012～2014 年にかけて，8.98%，11.41%，15.12% と変化しており，概ね同様の傾向である．ゴールドに対して，ブロンズの割合は減少している．世界と比較すると，過去の論文でのブロンズの割合が圧倒的に高いことがわかる．4.1 節では日本における OA の割合の高さの要因として「ブ

ロンズの割合の高さ」を挙げたが、これらは過去の論文に起因するものであることがわかる。

また、日本の論文件数に注目すると、2004年から2005年にかけて急激に上昇している。この原因としては、2002年9月にJSTリンクセンター[55]が設立されてCrossrefのサービスに参加したこと[56]によって、J-STAGEで公開されている論文を中心にCrossref DOIの付与が進んだことが原因だと思われる。図10にDOIが付与されていない論文も含めたScopusでの日本の論文件数の推移を示すが、図8の論文件数の推移と比較すると、2004年から2005年にかけての上昇幅が小さい。

ゴールドを細分化した詳細な結果については、出版年による内訳の変化が少ないことから省略する。

4.3.2 リポジトリによるOA

図11に、日本と世界でのリポジトリでOAとなっている論文の割合を出版年ごとに示す。日本と世界は概ね同様の傾向であるが、近年では日本のほうが若干高い。図12に日本の論文のみを対象として、リポジトリの種別などに注目したより詳細な結果を表す。機関リポジトリとその他リポジトリでは、いずれの出版年でもその他リポジトリで公開されている論文の割合が圧倒的に高い。これらのリポジトリとしては、Semantic Scholar, PubMed Central, Europe PMC, arXivが挙げられる(4.4節参照)。特に、2012~2016年にかけてその他リポジトリで公開されている論文の割合は増加している。図8では当該期間にグリーンの割合は増加しておらず、ゴールドの割合が顕著に伸びている。ゴールドで出版された論文の多くにはCC-BY等のライセンスが付与されていることから、その他リポジトリでこれらの論文の公開が進んだと考えられる。機関リポジトリで公開されている論文の割合は、2010年以降は5%程度である。また、機関リポジトリのみで公開されている論文の割合は1~2%程度である。この割合は近年は減少傾向にあり、理由の一つとして出版者の定めるエンバーゴ期間が挙げられる。また、OAに対する意識が高い著者がOAジャーナルでの出版を選択することで、相対的に機関リポジトリのみで公開される論文が減少してしまうことも一因かもしれない。図8では、近年はゴールドの割合が上昇していることがわかる。

4.4 OAに使用されているリポジトリ

リポジトリによるOAについて詳しく探るため、論文が公開されているリポジトリの調査を実施した。3.2.2項で述べたとおり、Unpaywallの各論文には、OAである場合、`oa_locations`というフィールドが存在する。`oa_locations`の値はリストであり、リストの各オブジェクトにはOAとなっているサイトについての情報が格納されている。その中でもフィールド`host_type`には`publisher`または`repository`という値が割り当てられており、「そのサイトが出版者であるかリポジトリであるか」ということがわかる。調査では各論文のフィールド`oa_locations`からオブジェクトのタイプが`repository`であるものを抽出し、`url`のネットワークロケーションとその頻度を集計した。

日本の論文を対象とした結果を表5に示す。公開している論文件数が最も高いリポジトリはSemantic Scholarである。Semantic Scholarは計算機科学分野と生物医学分野の論文の検索プラットフォームである。一般的には著者自身でリポジトリに登録することができず、あらゆるリポジトリから論文のPDFファイルを収集して公開するアグリゲータとして捉えられる。2位と3位は、PubMed CentralとEurope PMCである。PubMed CentralやEurope PMCは著者がアップロードすることも可能である[57, 58]が、多くの場合は出版者等からPDFを取得して公開している。著者のアップロードによる公開がメインとなっている機関リポジトリ以外のリポジトリとしては、arXiv(4位)、figshare(7位)、CERN Document Server(8位)が挙げられる。機関リポジトリとしては、京都大学(5位)、北海道大学(6位)、筑波大学(9位)、金沢大学(10位)のものが上位10件に含まれている。以上のことから、機関リポジトリを除いて日本の論文の公開に利用されているリポジトリとしては、arXiv, figshare, CERN Document Serverが利用頻度が高いといえる。なお、海外の共著者が論文をリポジトリに登録している可能性もあるため、日本の著者によるOAで利用されているリポジトリであるとは必ずしもいえない。

さらに、世界の論文を対象とした結果を表6に示す。1~4位は表5の日本の結果と同様である。5位はOpenAIREとCERNによって開発されたZenodoである。データセットを含めた様々な成果を共有で

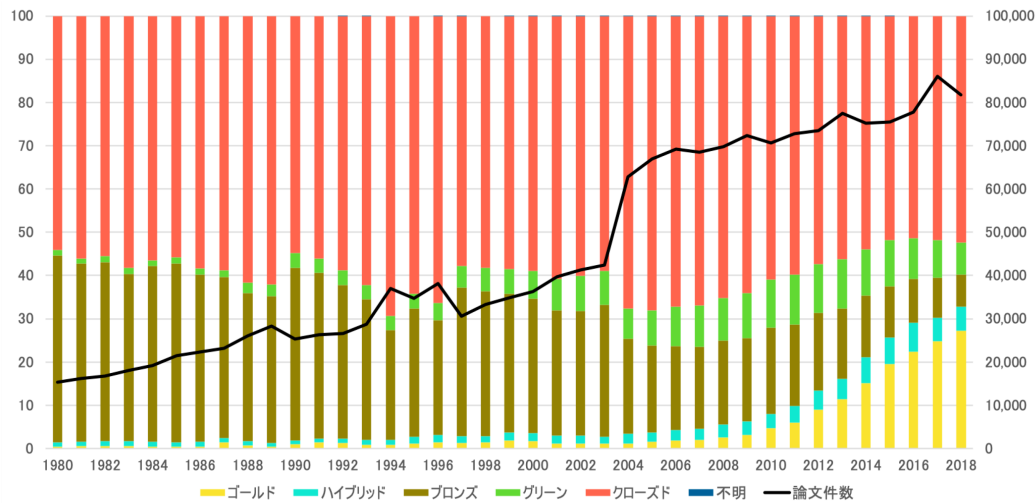


図 8: 日本における出版年ごとの OA 状況.

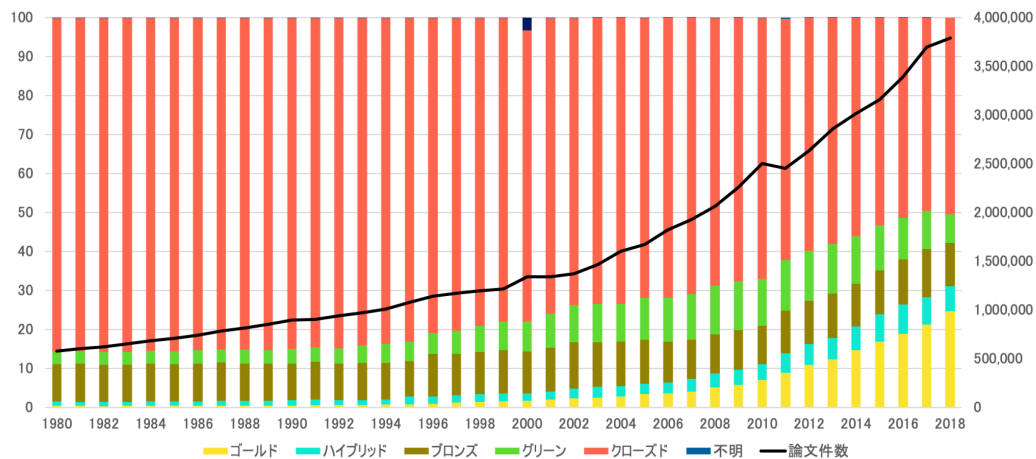


図 9: 世界における出版年ごとの OA 状況.

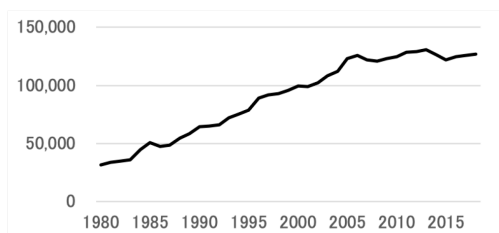


図 10: Scopus での日本の論文件数の推移.

きるリポジトリであり、論文も多く収録している [59, 60]. 6 位の Hyper Articles en Ligne (HAL) はフランス全土のリポジトリとして捉えられ、多くの論文が登録されている. 8 位は永久識別子を提供するハンドリングシステムであり、別にランディングペー

ジが存在すると考えられる. 9 位は DergiPark というトルコの学術雑誌のホスティングと編集管理システムを提供しているサービスであり、リポジトリではなく出版者であると考えられる. 詳しくみたところ、これらの論文の `oa_locations` には英語ページとトルコ語ページが格納されており、英語ページは `host_type` が `publisher` であるのに対して、トルコ語のものでは `repository` となっている. そのため、リポジトリとして捉えるのは誤りであると思われる. 10 位は HathiTrust の電子図書館である.

なお表 5 と表 6 の論文件数にカウントされているものは、出版者で OA となっているものも含まれており、全てが表 1 におけるグリーンではない. グリーンに貢献しているリポジトリ（出版者では OA

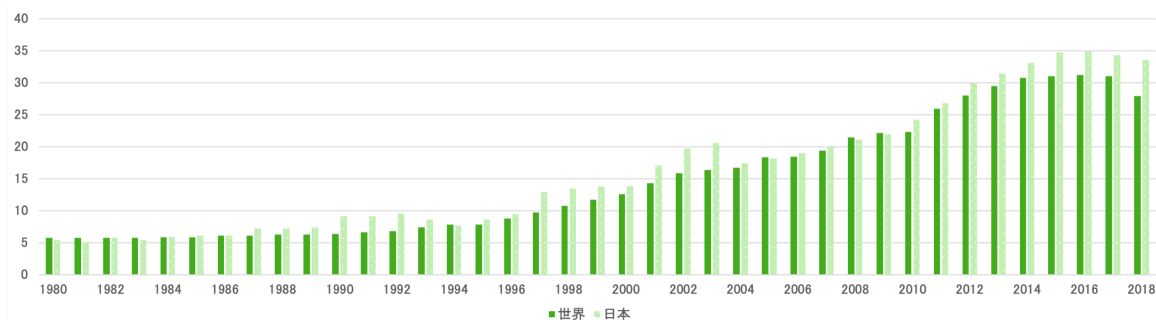


図 11: 出版年ごとのリポジトリによる OA 状況.

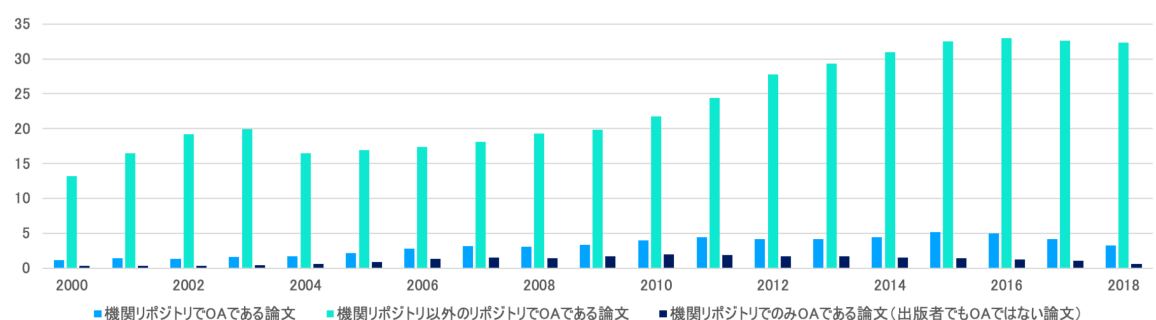


図 12: 日本における出版年ごとのリポジトリによる OA 状況.

でない論文を多く公開するリポジトリ)を調査したところ、概ね表 5 と表 6 の順位に変動はなかった。

表 5: 日本の論文が公開されているリポジトリ上位 10 件.

	リポジトリ	論文件数
1	pdfs.semanticscholar.org	272,902
2	www.ncbi.nlm.nih.gov	154,672
3	europemc.org	115,888
4	arxiv.org	28,875
5	repository.kulib.kyoto-u.ac.jp	11,937
6	eprints.lib.hokudai.ac.jp	9,322
7	figshare.com	5,038
8	cds.cern.ch	4,944
9	tsukuba.repo.nii.ac.jp	4,494
10	kanazawa-u.repo.nii.ac.jp	3,555

5 考察

本章では、4 章で得られた結果を考察する。

表 3 より、日本の OA である論文の割合は 41.83% であり、世界における割合 (29.77%) よりも高いことが判明した。日本の論文の OA の割合の高さの理

表 6: 世界の論文が公開されているリポジトリ上位 10 件.

	リポジトリ	論文件数
1	pdfs.semanticscholar.org	8,714,986
2	www.ncbi.nlm.nih.gov	4,549,458
3	europemc.org	3,830,596
4	arxiv.org	895,370
5	zenodo.org	339,266
6	hal.archives-ouvertes.fr	285,491
7	cds.cern.ch	162,320
8	hdl.handle.net	124,104
9	dergipark.org.tr	103,632
10	babel.hathitrust.org	101,256

由として、ブロンズの割合の高さが挙げられる。図 8 より、過去の論文によってブロンズの割合が高いことが判明した。特にその割合が高い 1980~2003 年に出版されたブロンズの 228,834 件の論文うち、189,272 件 (82.71%) が J-STAGE で公開されていることがわかった。1980~2003 年に出版された出版者で OA である論文 (ゴールド、ハイブリッド、ブロンズのいずれかの OA 種別が付与される論文) は 251,838 件あるが、そのうち J-STAGE で公開され

ているものは194,321件(77.16%)であった。よって、他の出版者と比較すると、J-STAGEにはライセンスが付与されていない過去の論文の割合が高いといえる。これらの論文について、著者から学協会に著作権が譲渡されている場合にはCC-BYライセンスを一括して適及的に付与することも法的には可能であるはずであるが、実際には広範な再利用を認めるライセンスを適及的に、著者の了承を得ず付与することには大きな反発が起こることも予想される。個別の著者に了承を得る労力も大きいことを考えると、当面は適及的なゴールドへの移行は困難であろう。

日本と世界共通の近年の傾向として、ゴールドの割合の上昇が挙げられる。4.2節では、学際的領域ではメガジャーナルの影響でゴールドの割合が高くなっていることを報告した。全体のゴールドの割合の上昇へのメガジャーナルの影響についても調査した。ここではメガジャーナルとしてScientific ReportsとPLOS ONEを考慮する。2014~2018年の各出版年において日本における前年からのゴールドの論文の増減率を計算したところ、それぞれ28.68%, 29.04%, 18.29%, 23.23%, 4.15%であった。対してメガジャーナルの増減率はそれぞれ2.50%, 21.99%, 11.34%, 7.32%, -15.65%であった。よって、日本のゴールドの割合の上昇はメガジャーナルではなく、その他のOAジャーナルによるものであるといえる。

リポジトリによるOAについての調査では、機関リポジトリで公開されている論文の割合は2.36%、機関リポジトリのみで公開されている論文の割合は0.83%であった。以前にも指摘されているとおり、日本の機関リポジトリのグリーンへの貢献は小さい^[48]。大学の機関リポジトリでは、一定の人員を割いて、機関リポジトリに登録されていない論文の著者にその論文の提出を促すメールを送付しているところもある^[61]。学術情報空間全体におけるOAの割合を増加させるためにも、OA情報を参照して、どこでもOAになっていない論文から優先的に教員に提出をお願いする、といった工夫が求められる。グリーンOAを認めているジャーナルに掲載されており、他の機関リポジトリで既に公開されている論文については、SWORD^[62]といったプロトコルを利用して機関リポジトリ間で論文の共有を行い、大学の貢献を形作る新しいチャンネルといった役割を果たしていくといったことが考えられる。多くの日本の論文を公開しているSemantic Scholar, PubMed

Centralといった外部のリポジトリとの連携も考えられるだろう。

このような工夫を可能にするには、各機関がそれぞれのOAの現状を把握できることを可能にすることが必要である。国立大学図書館協会オープンアクセス委員会が2016~2017年にかけて実施した調査^[39]では、全体の論文件数やゴールドの論文件数を「不明」や概数で回答した大学図書館が約半数であった。各機関が現状を認識して戦略を立てやすくするためにも、ドイツにおけるOpen Access Monitor^[26]、デンマークにおけるOpen Access Indicator^[63]のような、現状の把握と評価を継続的に可能にする仕組みやツールの開発が必要だろう。

Open Access Monitorは、有償・無償のデータベースからドイツの各機関の論文のデータを継続的に取得することによって構築されている。持続可能性の観点から、可能な限り無償のデータベースのみで構築することが望ましい。CrossrefやJaLCといったDOI登録機関が提供するメタデータで、各著者とORCIDといった研究者の識別子や著者の各所属機関とRORといった機関の識別子の紐付けが進むと、APIを使用することで継続的に各機関の論文を把握しやすくなる。さらにCrossref DOIが付与されている論文については、UnpaywallのAPIを使用することでOA状況を把握することが可能となる。特に、CrossrefとORCIDは連携していることから、各機関が各研究者のORCIDを把握しており、論文のメタデータでのORCIDの普及が進めば、各機関の論文を把握することは可能になる。2020年3月までに登録されたCrossrefのレコード全件^[64]を使用して調査したところ、2015年以降に出版された学術雑誌論文19,007,035件では、著者として延べ79,935,641名がメタデータに記録されているが、そのうちORCIDと紐づけられている著者は延べ5,874,678名(7.35%)であった。このことから、現時点では実現は難しいと思われる。論文のメタデータにおいてORCIDやRORといった識別子の普及が進むことが求められる。

本稿では、Unpaywallのデータベースのスナップショットを使用して調査を実施したが、全てのリポジトリが対象となっているわけではなく、日本の機関リポジトリは対象となっていないものも多い(3.2.1項参照)。IRDBで「本文あり」「資源タイプがjournal-article」「言語が英語」という3つの条件を満たす論文を検索したところ、106,519件

がヒットした（2020年2月26日時点）。これらの条件を満たす論文全件に Crossref DOI が付与されていることを想定すると、機関リポジトリで OA である論文の割合は 5.32% であるといえる。これらの論文全件が Scopus の収録対象となっているわけではないので実際の数値はわずかに小さいと考えられるが、本稿の調査での機関リポジトリで OA である論文の割合 2.36%（表 4 参照）と比較すると、2.25 倍の差異が存在することになる。このように実際の数値との乖離が懸念されるが、1 章で述べたとおり、Unpaywall は多くの研究者に使用されているプラグインである。よって機関リポジトリで OA となっている論文の発見可能性という観点から、Unpaywall を使用したこの調査には一定の意義があると考えられる。将来には、より正確なグリーンの状況を探るためにも、IRDB を使用した調査に取り組みたい。

6 おわりに

本稿は、Unpaywall のデータベースのスナップショットを利用して、日本における OA 状況の調査を実施した。日本の OA である論文の割合は 41.83% であり、世界における割合（29.77%）よりも高いことが判明した。この割合の高さの要因として、過去に出版された論文の多くがブロンズであることが挙げられる。日本と世界共通して、近年はゴールドの割合が上昇している。日本で論文の公開のために頻繁に使用されている機関リポジトリ以外のリポジトリとしては、arXiv, figshare, CERN Document Server であることが判明した。本稿の調査では、機関リポジトリに登録されている論文の割合は、近年は一定して 5% 程度であるとの結果が出たが、Unpaywall に把握されていない機関リポジトリも存在するため、IRDB を使用した調査も必要になってくる。

本稿では日本全体の結果のみを示したが、各機関が OA に関する取り組みの評価や戦略を策定するには、各機関ごとに OA 状況を継続的に把握する必要がある。欧州での取り組み^[26, 63]等を参考にしつつ、仕組みづくりが必要になってくるだろう。

謝辞

本研究は JSPS 科研費 18K13235, 20K20132, 稲盛研究助成, JST AIP JPMJCR19U1 の助成を受

けたものである。また本研究は、オープンアクセスリポジトリ推進協会（JPCOAR）コンテンツ流通促進作業部会でのプロジェクトの一つとして進められた。

参考文献

- [1] 佐藤翔: “オープンアクセスの広がりとは現在の争点”, 情報管理, Vol. 56, No. 7, pp. 414–424, 2013.
- [2] “Read the Budapest Open Access Initiative”, <https://www.budapestopenaccessinitiative.org/read>, 2002, 最終閲覧日: 2020 年 2 月 26 日.
- [3] 佐藤翔: “学術情報流通と図書館: オープンアクセスからオープンサイエンスへ”, 図書館界, Vol. 70, No. 1, pp. 245–264, 2018.
- [4] 時実象一: “オープンアクセス: 機関リポジトリの最近の動向”, 情報の科学と技術, Vol. 59, No. 5, pp. 231–237, 2009.
- [5] 時実象一: “オープンアクセスの動向 (1): オープンアクセスの義務化とその影響”, 情報の科学と技術, Vol. 64, No. 10, pp. 426–434, 2014.
- [6] 国立大学図書館協会: “オープンアクセスに関する声明”, https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/toushin/attach/1283016.htm, 2009, 最終閲覧日: 2020 年 3 月 6 日.
- [7] 内閣府: “第 4 期科学技術基本計画（平成 23～27 年度）”, <https://www8.cao.go.jp/cstp/kihonkeikaku/index4.html>, 2011, 最終閲覧日: 2020 年 3 月 6 日.
- [8] 国際的動向を踏まえたオープンサイエンスに関する検討会: “我が国におけるオープンサイエンス推進のあり方について～サイエンスの新たな飛躍の時代の幕開け～”, https://www8.cao.go.jp/cstp/sonota/openscience/150330_openscience_1.pdf, 2015, 最終閲覧日: 2020 年 3 月 6 日.

- [9] 森いづみ: “CA1753 – 動向レビュー: 大学キャンパスの中のオープンアクセス”, カレントアウェアネス, No. 309, pp. 14–17, 2011.
- [10] 京都大学図書館機構: “京都大学オープンアクセス方針”, <https://www.kulib.kyoto-u.ac.jp/content0/13092>, 最終閲覧日: 2020 年 3 月 7 日.
- [11] Piwowar, Heather; Priem, Jason; Larivière, Vincent; Alperin, Juan Pablo; Matthias, Lisa; Norlander, Bree; Farley, Ashley; West, Jevin; Haustein, Stefanie: “The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles”, PeerJ, Vol. 6, p. e4375, 2018.
- [12] Martín-Martín, Alberto; Costas, Rodrigo; van Leeuwen, Thed; López-Cózar, Emilio Delgado: “Evidence of open access of scientific publications in google scholar: A large-scale analysis”, Journal of Informetrics, Vol. 12, No. 3, pp. 819–841, 2018.
- [13] Piwowar, Heather; Priem, Jason; Orr, Richard: “The future of OA: A large-scale analysis projecting Open Access publication and readership”, bioRxiv, 2019.
- [14] Unpaywall: “Browser extension”, <https://unpaywall.org/products/extension>, 最終閲覧日: 2020 年 3 月 5 日.
- [15] Harnad, Stevan; Brody, Tim; Vallières, François; Carr, Les; Hitchcock, Steve; Gingras, Yves; Oppenheim, Charles; Stamerjohanns, Heinrich; Hilf, Eberhard R: “The access/impact problem and the green and gold roads to Open Access”, Serials review, Vol. 30, No. 4, pp. 310–314, 2004.
- [16] Harnad, Stevan; Brody, Tim; Vallières, François; Carr, Les; Hitchcock, Steve; Gingras, Yves; Oppenheim, Charles; Hajjem, Chawki; Hilf, Eberhard R: “The access/impact problem and the green and gold roads to Open Access: An update”, Serials review, Vol. 34, No. 1, pp. 36–40, 2008.
- [17] Suber, Peter: “Gratis and libre open access”, SPARC Open Access Newsletter, 2008.
- [18] Laakso, Mikael; Björk, Bo-Christer: “Delayed open access: An overlooked high-impact category of openly available scientific literature”, Journal of the American Society for Information Science and Technology, Vol. 64, No. 7, pp. 1323–1329, 2013.
- [19] “ResearchGate”, <https://www.researchgate.net/>, 最終閲覧日: 2020 年 2 月 14 日.
- [20] “Academia”, <https://www.academia.edu/>, 最終閲覧日: 2020 年 2 月 14 日.
- [21] Björk, Bo-Christer: “Gold, green, and black open access.”, Learned publishing, Vol. 30, No. 2, pp. 173–175, 2017.
- [22] Archambault, Éric; Amyot, Didier; Deschamps, Philippe; Nicol, Aurore; Provencher, Françoise; Rebout, Lise; Roberge, Guillaume: “Proportion of open access papers published in peer-reviewed journals at the European and world levels – 1996-2013”, RTD-B6-PP-2011-2: Study to develop a set of indicators to measure open access, 2014.
- [23] Corral, Sheila: “The open movement: What libraries can do.”, In Proceedings of the Charleston Library Conference, pp. 559–570, Purdue University Press, 2015.
- [24] Haschak, Paul G: “The ‘platinum route’ to open access: a case study of E-JASL: The Electronic journal of academic and special librarianship”, Information Research, Vol. 12, No. 4, 2007.
- [25] Watson, Roger: “Gold, green, or diamond: Which is best?”, Nurse Author & Editor, Vol. 28, No. 2, p. 6, 2018.
- [26] Forschungszentrum Jülich: “Open Access Monitor”, <https://open-access-monitor.de>, 最終閲覧日: 2020 年 2 月 18 日.

- [27] 尾鷲瑞穂; 野崎久美子; 張替香織; 村上章人: “国立環境研究所におけるオープンアクセス実態調査”, 情報の科学と技術, Vol. 68, No. 10, pp. 500–505, 2018.
- [28] Jamali, Hamid R: “Copyright compliance and infringement in ResearchGate full-text journal articles”, Scientometrics, Vol. 112, No. 1, pp. 241–254, 2017.
- [29] Fortney, Katie; Gonder, Justin: “A social networking site is not an open access repository”, <https://osc.universityofcalifornia.edu/2015/12/a-social-networking-site-is-not-an-open-access-repository/>, 2015, 最終閲覧日: 2020 年 2 月 14 日.
- [30] Björk, Bo-Christer: “The open access movement at a crossroad: Are the big publishers and academic social media taking over?”, Learned Publishing, Vol. 29, No. 2, pp. 131–134, 2016.
- [31] Shamash, Katie: “Article processing charges (APCs) and subscriptions”, <https://www.jisc.ac.uk/reports/apcs-and-subscriptions>, 2002, 最終閲覧日: 2020 年 3 月 5 日.
- [32] “Directory of Open Access Journals (DOAJ)”, <https://doaj.org/>, 最終閲覧日: 2019 年 11 月 29 日.
- [33] Beall, Jeffrey: “Predatory publishers are corrupting open access”, Nature, Vol. 489, No. 7415, p. 179, 2012.
- [34] 千葉浩之: “CA1960 – ハゲタカジャーナル問題: 大学図書館員の視点から”, カレントアウェアネス, , No. 341, pp. 12–14, 2019.
- [35] Lynch, Clifford A: “Institutional repositories: Essential infrastructure for scholarship in the digital age”, portal: Libraries and the Academy, Vol. 3, No. 2, pp. 327–336, 2003.
- [36] “機関リポジトリ: デジタル時代における学術研究に不可欠のインフラストラクチャ”, <https://www.nii.ac.jp/irp/archive/translation/ar1/>, 最終閲覧日: 2020 年 2 月 20 日.
- [37] 福澤尚美: “ジャーナルに注目した主要国の論文発表の特徴: オープンアクセス、出版国、使用言語の分析”, In 研究・イノベーション学会年次学術大会講演要旨集, pp. 624–627, 研究・イノベーション学会, 2016.
- [38] 池内有為; 林和弘; 赤池伸一: “研究データ公開と論文のオープンアクセスに関する実態調査”, NISTEP RESEARCH MATERIAL, , No. 268, pp. 1–108, 2017.
- [39] 国立大学図書館協会オープンアクセス委員会: “「オープンアクセスへの取り組み状況に関する実態調査」報告書”, https://www.janul.jp/j/projects/oa/OA_report_201703.pdf, 2017, 最終閲覧日: 2020 年 2 月 26 日.
- [40] 小陳左和子; 矢野恵子: “ジャーナル購読からオープンアクセス出版への転換に向けて”, 大学図書館研究, Vol. 109, p. 2015, 2018.
- [41] 尾城孝一; 市古みどり: “オープンアクセスの現在地とその先にあるもの”, 大学図書館研究, Vol. 109, p. 2014, 2018.
- [42] “Affiliation Retrieval API”, <https://dev.elsevier.com/documentation/AffiliationRetrievalAPI.wadl>, 2019, 最終閲覧日: 2019 年 12 月 29 日.
- [43] “Journal List”, <http://help.incites.clarivate.com/incitesLiveESI/ESIGroup/overviewESI/esiJournalsList.html>, 2019, 最終閲覧日: 2019 年 12 月 18 日.
- [44] Schuldt, Karsten: “On Open Access Analytics in Europe”, <https://sonar.ch/post/open-access-analytics-europe/>, 2019, 最終閲覧日: 2020 年 2 月 18 日.
- [45] Priem, Jason: “What do the types of oa_status (green, gold, hybrid, and bronze) mean?”, <https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and->

- bronze-mean-, 2019, 最終閲覧日: 2020 年 2 月 18 日.
- [46] Priem, Jason: “How do we decide if a given journal is fully OA?”, <https://support.unpaywall.org/support/solutions/articles/44001792752-how-do-we-decide-if-a-given-journal-is-fully-oa->, 2019, 最終閲覧日: 2020 年 2 月 18 日.
- [47] “学術機関リポジトリデータベース (IRDB)”, <https://irdb.nii.ac.jp/>, 最終閲覧日: 2020 年 2 月 26 日.
- [48] 国公立大学図書館協力委員会シンポジウム企画・運営委員会: “平成 28 年度大学図書館シンポジウム「10 年後の大学図書館を考える: オープンアクセス時代の大学図書館の新たな役割」報告”, 大学図書館研究, Vol. 110, p. 2025, 2018.
- [49] Chan, Leslie: “Reimagining the role of institutional repositories in open scholarship”, https://www.coar-repositories.org/files/2_OpenAIRE-COAR_keynotes_LeslieChan.pdf, 最終閲覧日: 2020 年 2 月 20 日.
- [50] “OA Location object”, <https://unpaywall.org/data-format#oa-location-object>, 最終閲覧日: 2020 年 2 月 18 日.
- [51] Python Software Foundation: “url-lib.parse”, <https://docs.python.org/ja/3/library/urllib.parse.html#module-urllib.parse>, 最終閲覧日: 2020 年 2 月 18 日.
- [52] Directory of Open Access Journals (DOAJ): “Guide to applying”, <https://doaj.org/apply/guide/>, 最終閲覧日: 2020 年 2 月 18 日.
- [53] Elsevier: “Content Policy and Selection”, <https://www.elsevier.com/solutions/scopus/how-scopus-works/content/content-policy-and-selection>, 最終閲覧日: 2020 年 2 月 25 日.
- [54] 佐藤翔: “オープンアクセスメジャーナルの興隆, と, 停滞”, 情報の科学と技術, Vol. 68, No. 4, pp. 187–188, 2018.
- [55] 白木澤佳子; 水野充: “JST リンクセンター”, 情報管理, Vol. 45, No. 7, pp. 502–505, 2002.
- [56] 独立行政法人科学技術振興機構: “J-STAGE 500 誌へのあゆみ”, J-STAGE NEWS, Vol. 特別号, p. 3, 2008.
- [57] National Institutes of Health (NIH): “PubMed Central Submission Assistance”, <https://www.nihlibrary.nih.gov/services/editing/pubmed-central-submission-assistance>, 最終閲覧日: 2020 年 2 月 25 日.
- [58] Europe PMC plus: “User guide”, <https://plus.europepmc.org/user-guide>, 最終閲覧日: 2020 年 2 月 25 日.
- [59] Sicilia, Miguel-Angel; García-Barriocanal, Elena; Sánchez-Alonso, Salvador: “Community curation in open dataset repositories: Insights from Zenodo”, Procedia Computer Science, Vol. 106, pp. 54–60, 2017.
- [60] Peters, Isabella; Kraker, Peter; Lex, Elisabeth; Gumpenberger, Christian; Gorraiz, Juan Ignacio: “Zenodo in the spotlight of traditional and new metrics”, Frontiers in Research Metrics and Analytics, Vol. 2, p. 13, 2017.
- [61] 河合将志; 林正治; 尾城孝一; 新妻聡; 西澤正己; 山地一禎: “グリーンオープンアクセス進捗度に関する計量分析”, 情報知識学会誌, Vol. 28, No. 4, pp. 298–301, 2018.
- [62] “SWORD”, <http://swordapp.org/>, 最終閲覧日: 2020 年 11 月 28 日.
- [63] Ministry of Higher Education and Science: “Danish Open Access Indicator”, <https://www.oaindikator.dk/en>, 最終閲覧日: 2020 年 3 月 9 日.
- [64] “March 2020 Public Data File from Crossref”, <http://dx.doi.org/10.13003/83b2gp>, 最終閲覧日: 2020 年 11 月 28 日.

(2020年 5月15日 受付)

(2021年 1月19日 採択)

(2020年 3月25日 プレプリント公開)

<http://hdl.handle.net/2433/246424>