

学術情報のオープン化の推進について（中間まとめ案）

1 検討の背景

- 近年、情報通信技術の急速な進展に伴い、研究成果（論文、生成された研究データ等）を分野や国境を越えて活用し、新たな価値を生み出すための取組が世界的に広まりつつある。研究成果のオープン化は、研究成果の相互利用を促進し、知の創出に新たな道を開くことが期待されるが、とりわけ研究データ等をオープン化しデータ駆動型の研究を推進することでイノベーションの創出につなげることを目指した新たな科学の進め方が注目されている。すなわち、研究成果のオープン化に関する議論は、従来の論文へのアクセスを中心としたオープンアクセスの概念にとどまらず、研究データを含む研究成果の利活用へと概念が変化し、研究の進め方の変化や新たな手法が生じつつあることを示している。

- 特に、諸外国では、公的な研究資金を用いた研究成果について、研究者コミュニティはもとより広く社会からのアクセスや利用を可能にするオープンアクセス、オープンデータの取組が顕著となっている¹。

- 我が国においては、論文のオープンアクセスについては、大学を中心として機関リポジトリによる公開が推進され、その取組が広まってきているが、研究成果の利活用を促進する観点からさらに推進する必要がある。また、研究データについては、研究分野によりオープン化に対する取組に違いがあるが、研究者コミュニティ及び関係機関においては、その意義を共有し、国際的な動向も踏まえ着実に取組を進めることが求められる。
研究成果のオープン化によって、我が国の学術研究等に新しい研究方法の拡大など新たな展開をもたらすことも課題である。

¹ 例えば、米国では、国立衛生研究所（NIH）が、助成を受けた査読論文の指定リポジトリへの登載を2008年に義務化。2013年には、科学技術政策局（OSTP）の指令により、NIH及び米国立科学財団（NSF）等がパブリックアクセスプランを策定し、査読論文及び研究データを指定するリポジトリで公開することを進めている（NIH及びNSFは、以前より研究データの公開について、それぞれのポリシーに沿った取組を行っている）。英国では、英国研究会議（RCUK）が、論文についてオープンアクセスジャーナルへ投稿することを内容とする新たなオープンアクセスポリシーを2012年に公開している。ドイツでは、ドイツ研究振興協会（DFG）が、オープンアクセスジャーナルに投稿した際の投稿料を、研究者の所属機関に助成することにより、オープンアクセスを推進している。なお、国際的な組織では、国際科学会議（ICSU）の委員会（ICSU-WDS, World Data System）が、科学データに関する国際的な取組の高度化を目的に活動しているほか、米国のNSFやEU等により設立された団体であるRDA（Research Data Alliance）が、研究データ流通のルール策定を目的に活動している。

- 本委員会は、このような認識の下、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」による報告²等を参考としつつ、学術情報³のオープン化に係る基本的方策について検討を進め、その要点を以下のとおりまとめた。

2 基本的考え方

- 新たな知を創出する学術研究等の成果は、人類社会の持続的発展の基礎となる共通の知的資産として共有されることが望ましいことから、大学及び研究機関（以下、「大学等」）における研究成果は原則公開⁴し、広く利活用されることを、研究者等が基本理念として共有する必要がある。

- 研究成果の公開を通じた利活用を促進することにより、自然科学のみならず、人文学・社会科学を含め、分野を越えた新たな知見の創出⁵や効率的な研究の推進等に資するとともに、研究成果への理解促進や研究成果のさらなる普及につながることを期待される。

なお、研究成果の利活用を促進する観点とは異なるが、研究の透明性を確保することや研究の過度な重複を避けることによって研究費を効率的に活用する観点からもオープン化の取組が求められる。

また、公的研究資金による研究成果は、広く社会に還元すべきものであることに鑑み、そのオープン化推進の必要性はなお一層強い。

- これらの意義を踏まえ、公的研究資金による研究成果のうち、論文及び論文のエビデンスとしての研究データ⁶は原則公開とすべきである。

- 論文のエビデンスとしての研究データの公開及び利活用を促進する前提として、まずデータが、研究者において適切に保管されることが重要である⁷。その上

² 「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」（平成 27 年 3 月 30 日 国際動向を踏まえたオープンサイエンスに関する検討会）。以下、「内閣府検討会報告書」と表記する。

³ 研究者や学生が研究・学習を進める上で必要な論文や各種データ等をいう。

⁴ 研究の成果としての論文や研究データを、インターネット上で公表し、合法的な用途で利用することを障壁無しで許可することを意味する。以下の記述において「公開」とある場合は、同じ意味で用いている。

⁵ 例えば、日本語による成果の公開及び利活用を通して、さらに学際的な研究が進展する可能性があることや情報発信の強化につながることも期待される。

⁶ 論文の裏付けとなるデータなど、研究結果を立証するために必要な科学界で共通に受け入れられるデジタル的に記録された事実に基づくデータ。

⁷ なお、人文学・社会科学の分野においては、一次史料等を電子化したものがデータとして扱われることが考えられるが、このような史料等への永続的なアクセスを保証するものとして、アーカイブ

で、保管されたデータについて、どのデータをどのような様式で研究データとして公開とすべきか、あるいはどのような場合に非公開とすべきかということについては、研究者コミュニティ等による検討を踏まえた対応が必要である。

3 研究成果の公開についての基本的方策

(1) 論文のオープンアクセスについての取組

○ オープンアクセスを推進する意義は、論文への自由なアクセスを保障するのみならず、利活用を促進することで、研究開発の費用対効果を上げるとともに、学際的な研究を促し、イノベーションの創出や科学そのものの発展を期待することである。

このため、これまで我が国において推進してきた取組を踏まえ、公的研究資金による論文については、原則公開とすることを第5期科学技術基本計画期間中に実行すべきである⁸。

○ オープンアクセスを推進する方策については、機関リポジトリ⁹をセルフアーカイブ（グリーンOA）の基盤として拡充するとともに、オープンアクセスジャーナルの育成を図っていく必要がある。

【研究資金配分機関が行うべき取組】

・論文のオープンアクセスに係る実施方針を定め、研究者等への周知を含めて計画的に取り組む。

【大学等に期待される取組】

・機関リポジトリをグリーンOAの基盤としてさらに拡充する。
・オープンアクセスに係る方針を定め公表することが望ましい。

が重要となる。

⁸ 2013年に開催されたG8科学技術大臣及びアカデミー会長会合等において、各国において研究成果のオープンアクセスを拡大させる方針が確認されるなど、国際的な取組が加速している。

⁹ 大学等における機関リポジトリの役割について、情報発信の重要な手段としてのみならず、次のように位置づけられている。

①大学等の生産する知的情報・資料の集積、長期保存の場（アーカイブ）

②学術情報の発信及び流通の基盤（論文、データ、報告書等の公表及び提供）

③学習・教育のための基盤（教材の電子化、提供、保存）

（「学術情報の国際発信・流通力強化に向けた基盤整備の充実について」（平成24年7月 科学・技術学術審議会学術分科会研究環境基盤部会学術情報基盤作業部会）より抜粋）

【科学技術振興機構（JST）が行うべき取組】

- ・我が国の公的支援による出版プラットフォームである J-STAGE について、レビュー誌の発信などを通じて国際的な存在感の向上を図る。

【国立情報学研究所（NII）が行うべき取組】

- ・機関リポジトリ構築の共用プラットフォーム（JAIRO Cloud）により、大学等における効率的な整備を支援する。
- ・国際学術情報流通基盤整備事業（SPARC Japan）によりセミナーを開催するなど、オープンアクセスに対する理解増進を図る。

【学協会に期待される取組】

- ・学協会が共同して質の高いオープンアクセスジャーナルを構築する。

【国が行うべき支援】

- ・JST が運用する J-STAGE について、レビュー誌の構築や利便性の高いインターフェイスの構築など、その高度化を支援する。
- ・NII が行うオープンアクセスに係る事業を支援する。
- ・科学研究費助成事業（研究成果公開促進費）により、オープンアクセスジャーナルの刊行を支援する。

（2）論文のエビデンスとしての研究データの公開

①分野ごとに異なる特性

- 研究分野によつての取組の状況は多様である。

例えば、天文学、素粒子物理学の分野のように、研究データの共有¹⁰がその研究者コミュニティにおいて既に標準の取組となっている分野もある。また、ライフサイエンス分野においては、先駆的に統合データベースが整備されるなど以前から研究データの共有の取組が行われており、今後も他分野の先例となる取組の促進が期待される¹¹。材料科学分野においては、マテリアルズインフォマティクスの重要性が認識され、研究データの共有の取組が始まっている。また、人文学・社会科学分野においては、研究データの捉え方が一様ではない

¹⁰ 研究データを研究者コミュニティの中で相互利用することを意味する。以下の記述において「共有」とある場合は、同じ意味で用いている。

¹¹ 研究データ共有の取組が進められる一方で、課題も顕在化してきている。具体的には、メタデータを付与して保管するだけでは利活用されにくいこと、データフォーマットや用語の標準化、研究データの質の保証、公開及び共有の義務化ルールの設定、公開されない研究データが多いこと、研究データの権利関係の明確化、研究データ公開のタイミング、プロジェクト終了後の研究データの維持、データサイエンティスト等の人材不足などが、委員会の審議の中で指摘された。

が、古典籍や文書など研究素材としての一次資料の電子化は着実に進展しており、多様なデジタルアーカイブが存在する。

②研究データの保管・管理

- 研究データの保管・管理は研究データの公開を進めるための前提であり、特に、公的研究資金による研究プロジェクト等の終了後も、研究データが利活用可能な状態で適切に管理されるよう、プロジェクト等の規模（例えば、一定額以上の規模に係るものを対象とする）やその目的及び分野の特性等に応じ、データ管理計画を作成し計画に従った管理を行うことが必要となる。

【研究資金配分機関が行うべき取組】

- ・申請者に、必要に応じデータ管理計画の提出を求める。

【大学等に期待される取組】

- ・研究者のデータ管理計画の作成と計画に従った管理の実施について支援する。
- ・研究データの保管に係る基盤を整備するに当たって、情報基盤の共有や効率的な整備の観点から、アカデミッククラウドの活用を図る。

【NIIが行うべき取組】

- ・アカデミッククラウドの構築に当たり、フォーマットの標準化やシステム開発及び共同調達等について、大学等と連携し進める。

【国が行うべき支援】

- ・NIIと大学等が連携して進めるアカデミッククラウドの構築を支援する。

③公開の対象とする研究データの範囲とその様式

- 公開すべき研究データ¹²の範囲とその様式については、国際的な動向や原則公開とする趣旨を踏まえた上で、学協会等の場において、研究上の必要性等を考慮した検討を行い、日本学術会議で研究者コミュニティのコンセンサスを形成していくことが求められる¹³。また、研究データは利活用可能な様式で公開されること及び信頼性が確保されていることが重要であり、研究データを基に

¹² 公開される研究データは、利活用可能な形式でなければならない。また、メタデータ、数値データ、テキスト、イメージ、ビジュアルデータなど多様なデータがあり、データを扱うプログラムがある場合はこれも含まれる。

¹³ 日本学術会議においては、オープンサイエンスの取組に関する検討委員会において、我が国科学界がとるべきオープンサイエンスへの対応について調査審議を行っている。

した精度の高い成果の再現を可能とするためには、研究で使用したプログラムやソフトウェアのバージョン等についても、データ作成者が公開時に明示する必要がある。

- 公的研究資金による研究成果である論文のエビデンスとしての研究データは原則公開とするが、機密保持等の観点から公開に制限がある場合などは、公開適用対象外とする¹⁴。

また、研究データへのアクセスや利用には、個人のプライバシー保護、財産的価値のある成果物の保護の観点から、制限事項を設ける必要がある。

【学協会等及び日本学術会議に求められる取組】

- ・学協会等は、研究上の必要性等を考慮した検討を行い、日本学術会議は、学協会等の意見を取りまとめ、研究者コミュニティとしてのコンセンサスを形成していくことが求められる。

【研究資金配分機関が行うべき取組】

- ・研究者コミュニティのコンセンサスを踏まえつつ、公開の進め方や管理方策及び非公開とする研究データや公開の様式等について公募要領等で示し、研究データの公開について推奨する¹⁵。

④研究データの公開方法

- 研究データの公開は、既に分野別の公的なデータベース¹⁶や学協会等で整備されているリポジトリがある場合、これらへの登載を促進することが妥当である。なお、公的なデータベース等がない分野については、研究成果の発信及び流通の基盤としての大学等のリポジトリを活用することが望ましい。

¹⁴ 内閣府検討会報告書では、「研究成果のうち、機密保持、企業秘密、国益及び国家安全保障に関わるもの、研究成果の商用化・産業化を目的として収集されたデータ、又は民間企業が保有するデータ並びに共同研究契約などで研究成果の公開に制限がある場合などは、公開適用対象外として考えるべきである。」としている。

¹⁵ 国の支援により統合データベースを整備しているライフサイエンス等の分野については、研究データの集積と利活用を促進する観点から、研究データの公開について一層推奨する必要がある。また、例えば、推奨に当たっては、研究データ共有の必要性が高い分野から推奨していくなど、実効性を考慮した対応が必要である。

¹⁶ 公的なデータベースの例としては、日本 DNA データバンク（DDBJ）、NBDC ヒトデータベースなどがある。

【研究資金配分機関が行うべき取組】

- ・ 研究データ公開の取組について公募要領等で推奨する。

【JST が行うべき取組】

- ・ J-STAGE の高度化と関連し、学協会の要望を踏まえ、データジャーナル出版に係る基盤の整備を行う。

【NII が行うべき取組】

- ・ JST 等と連携して、各データベースや各機関のリポジトリ等に搭載されている研究データの横断的な検索・利活用を可能とするための基盤の整備を行い、サービスを提供する。

【学協会に期待される取組】

- ・ 複数の学協会が共同して、データジャーナルを構築する。

【国が行うべき支援】

- ・ JST 及び NII が行う基盤整備を支援する。
- ・ 科学研究費助成事業（研究成果公開促進費）により、データジャーナルの刊行を支援する。

(3) 研究成果の散逸等の防止

大学等においては、研究成果を知的資産と捉え、明確な方針の下で保管し、蓄積していくことが重要である。

【大学等に期待される取組】

- ・ 論文、研究データの管理に係る規則¹⁷を定め、研究成果の散逸、消滅、損壊を防止するための施策を講ずる。
- ・ 具体的には、論文及び研究データに永続性のあるデジタル識別子を付与し管理する仕組みを確立する必要がある、ジャパンリンクセンター（JaLC）の活動と連携し進めることが望ましい。

【学協会に期待される取組】

- ・ 学協会が研究成果の著作権を譲り受けている場合には、著作権ポリシーを整備し、大学等の研究成果散逸等防止のための取組に協力する内容とすることが望ましい。

¹⁷ 規則には、大学等における研究成果の保管期間・方法について規定する必要がある。また、研究室等で個別に保管されている利活用可能な研究データについて、研究者の異動や退職に当たっての扱いについても検討しておく必要がある。

(4) 研究成果の利活用

①利用ルール等の明示

- 公開される研究成果の利活用を促進するためには、機関リポジトリ等を通じたオープンアクセスの取組を一層促進するとともに、別途著作権処理が必要となる利用についても、利用者が権利処理に負担を感じないで利活用できる仕組みを設けることが重要である。
- 論文については、学協会は、その刊行する学術誌に掲載される論文について、著作権の帰属や利用条件などの著作権ポリシーを明示する必要がある。自然科学分野等で一般的なように、学協会が研究者から著作権を譲り受ける場合には、著作権ポリシーにおいて利活用を促進するような利用条件を設けるとともに、利用者が簡便に権利処理をできるライセンス体制を構築することが期待される。
- 人文学・社会科学分野で見られるように、学協会が研究者から論文の著作権を譲り受けず、研究者が著作権を保有する場合にも、利用者が簡便に著作権処理できるようにすることが望ましい。そのために、研究者が、クリエイティブ・コモンズ・ライセンス¹⁸で利用条件を予め明示しておくことも考えられる。
- 研究データについても、利用ルールを明示し利活用を円滑化することが望ましい¹⁹。

【学協会が行うべき取組】

- ・刊行する学術誌に掲載される論文の著作権ポリシーを策定し、明示する。

¹⁸ クリエイティブ・コモンズ・ライセンス（CCライセンス）は、インターネット時代のための新しい著作権ルールの普及を目指し、CCライセンスを提供している国際的非営利組織とそのプロジェクトの総称。様々な作品の作者が、著作権に関わる利用条件について簡便に意思表示をするためのツールとして活用している。詳細は、<http://creativecommons.jp/licenses/>参照。研究成果についてCCライセンスを用いる場合、学協会が研究者から著作権を譲り受けていない論文については、著作権・著作権者等の出典を表示すれば自由な利用を許諾するCC-BYライセンスを採用することが考えられる。

¹⁹ 研究データは著作物でないとされるのが通例であるので、CCライセンスにおいて、著作権法上の何の制約もなく自由に利用できることを意味するCC0を採用するのが本来の姿であるが、研究データ作成への研究者のインセンティブを確保するためには、利活用の際にデータ作成者の表示を求めることは重要である。そのための便法として、CC-BYライセンスを利用してデータ作成者の表示を希望することを示すか、著作権法の規律とは別に、研究者コミュニティにおけるルール・慣行としてデータ作成者名の表示を求めることも考えられる。

②研究データの引用と評価の取組

- アクセス可能となった研究データの利用者は、論文などの引用と同じく引用元を明らかにする義務がある。この引用により、データ作成者の貢献が記録され、業績として評価することを、大学等及び研究者コミュニティにおいて共通に認識し、実行していく必要がある。
- 研究成果の利活用を促進する観点から、論文及び研究データに永続性のあるデジタル識別子を付与する必要があるが、どの粒度の研究データに対して識別子を付与すべきかについては、研究者コミュニティにおけるコンセンサスの形成が必要であり、これに取り組む JaLC の活動を推進する必要がある。

【大学等に期待される取組】

- ・ 引用されたデータ作成者の貢献を業績として評価する。

【国が行うべき支援】

- ・ デジタル識別子の付与に取り組む JaLC の活動を支援する。

(5) 人材育成の取組

研究データの公開を促進するためには、研究者の支援に係る人材や研究データを適切に取り扱える人材の育成が必要である。

【大学等に期待される取組】

- ・ 大学等においては、技術職員、URA 及び大学図書館職員等を中心としたデータ管理体制を構築し、研究者への支援に資するとともに、必要に応じて複数の大学等が共同して、データサイエンティストやデータキュレーターを育成するシステムを検討し、推進する。
- ・ 特に、大学図書館については、機関リポジトリの構築を進めてきた経験等から、研究成果の利活用促進の取組に積極的な役割を果たすことが期待される。このため、大学の当該領域に関連する研究科等において、大学図書館職員等を対象にデータキュレーターを育成するプログラムを開発し、実践的に取り組んでいく。

【国が行うべき支援】

- ・ 複数の大学等が共同して行うデータサイエンティスト等の人材を育成する取組を支援する。

4 研究データ基盤整備の方向性

- 研究データ基盤の整備に当たっては、国は、研究データが集積することの重要性に鑑み、研究データを的確に保存し活用していくためのプラットフォームの整備に重点を置くべきである。

- また、基盤の整備については、国際的な協調を図っていく視点も重要である。例えば、ジャーナルに論文の掲載が決定した際に、エビデンスとしての研究データの公開を求められる事例があるが、その研究データの公開先として我が国のデータベース等がプラットフォームの役割を果たせるようにしていく必要がある。国は、このような環境の構築に係る研究者コミュニティの活動を支援していく必要がある。

【国が行うべき支援】

- ・ 現在、地球環境情報統融合プログラム（DIAS-P）、バイオサイエンスデータセンター（NBDC）、センター・オブ・イノベーション（COI）プログラムなどの事業が行われており、研究データの蓄積や利活用が促進されている。プラットフォームの整備に関連しては、これら事業で蓄積された大量の研究データについて、分野を越え統合的な解析を実施することで、新たな価値を創造することが期待されており、このような新たな取組を国が支援していくことが重要である。

5 今後の検討について

日本学術会議及び内閣府のオープンサイエンス推進に関するフォローアップ検討会等において、オープンサイエンスに係る検討が進められていることから、その検討内容等も踏まえ、より効果的な推進方策等について検討していく必要がある。

用語解説

アーカイブ

本中間まとめでは、電子的な文書等を恒久的に保存する機能をもつ計算機システム、またはそこに収められている電子的ファイルをいう。

アカデミッククラウド

大学等が有する研究や教育に関するデータを大学等の間で共有するためのクラウド。

永続性のあるデジタル識別子

研究成果に対して、それを一義的に識別するデジタル情報であって、当該研究成果の有用性寿命と同等もしくはそれ以上永続するもの。

オープンアクセス

論文等の学術情報をインターネットから無料で入手でき、誰でも制約なくアクセスできるようにすること。オープンアクセスを実現する手段は多様であるが、機関リポジトリ、専門分野別のアーカイブなどへ研究者自らが論文等を掲載していくものと、購読料に依存しないビジネスモデルを備えた学術雑誌の刊行主体が行うものとに大別される。

オープンデータ

機械判読可能な研究データを、二次利用可能な研究データとして公開すること。

機関リポジトリ

大学等の機関において生産された電子的な知的生産物の保存や発信を行うためのインターネット上のアーカイブシステム。

グリーン OA

査読付き論文について出版社版または出版社版に至る前の著者最終原稿を大学等が構築・運用する機関リポジトリに登載し、公開する方法。グリーン OA の利点は、著者の負担なしに有料雑誌の論文情報が公開されることである。

研究資金配分機関

公的研究資金の配分を行う機関。

公的研究資金

文部科学省又は文部科学省が所管する独立行政法人から配分される競争的研究費。

ジャパンリンクセンター(JaLC)

電子化された論文等の学術コンテンツに、国際標準の識別番号(Digital Object Identifier, DOI)を付与する権限を持つ機関。科学技術振興機構や国立情報学研究所等が共同で運営している。

データ管理計画

研究プロジェクト等における研究データの取り扱いを定めるものであり、具体的にはデータの種類、フォーマット、アクセス及び共有のための方針、研究成果の保管に関する計画などについて記載されるもの。

データキュレーター

多くのデータの中から、利活用可能なデータを選び出し、それらを必要に応じ修復し、また、データを組み合わせることも含めて分析する役割を担う人材。

データサイエンティスト

ビックデータ処理技術、データ可視化技術、データ解析法を取得し、さらに研究戦略立案能力、コミュニケーション能力、研究倫理などのデータリテラシーを備えた人材。

プラットフォーム

コンピュータの動作を支える基盤(ハードウェアやOS)から、何らかの活動やサービスを実現する基盤まで幅広く利用される用語であるが、この中間まとめでは、研究データを集積、保存、発信するためのシステムを指す。

マテリアルズインフォマティクス

データ科学、計算科学と物質・材料の物理的・化学的性質に関する多様で膨大なデータとを駆使して、物質・材料科学の諸問題を解明するための科学技術的手法。

メタデータ

公開するデータ自体がどのようなデータであることを示す索引情報のこと。データの作成日時や作成者、データ形式、タイトル、注釈などが考えられる。データを一元的、かつ効率的に管理し、検索したりするために重要な情報である。

リサーチ・アドミニストレーター(URA)

大学等において、研究者とともに、研究企画立案、研究資金の調達・管理、知財の管理・活用等を行う人材。

レビュー誌

特定の主題分野やテーマに関して、既に出版された多数の原著論文を対象に、要約、解説、評価を行い、その分野(テーマ)の研究の現状や将来性等についてまとめた論文をレビュー(総説)といい、主としてそのような論文を掲載している雑誌。

J-STAGE

科学技術振興機構(JST)が構築・運営している学協会の電子ジャーナル出版支援及び公開のシステム。