

(2) オープンサイエンスのためのデータ基盤の整備

平成30年6月20日(水)

○目指すべき将来像

- ・国益や研究分野の特性等を踏まえて、オープン・アンド・クローズ戦略³⁵を考慮し、サイバー空間上での研究データ³⁶の保存・管理に取り組み、諸外国の研究データ基盤とも連携して巨大な「知の源泉」を構築し、あらゆる者が研究成果を幅広く活用
- ・その結果、所属機関、専門分野、国境を越えた新たな協働による知の創出が加速

○目標

<リポジトリ³⁷の整備及び展開>

- ・機関リポジトリ³⁸を活用した研究データの管理・公開・検索を促進するシステムを開発し、2020年度に運用開始

<研究データの管理・利活用についての方針・計画の策定等>

- ・研究成果としての研究データの管理・利活用のための方針・計画の策定³⁹を促進
- ・これらの方針・計画に基づき公的資金による研究データについて、機関リポジトリを始めとするデータインフラで公開を促進
- ・公的資金による研究成果としての研究データについては、データインフラを通して機械判読可能性と相互運用性を確保するとともに、公開する研究データについては諸外国の研究データ基盤との連携を促進

<人材の育成及び研究データ利活用の実態把握>

- ・研究データの利活用を図るため、研修教材の活用を促進するとともに、実態把握を行いながら、研究者や研究支援職員の意識を向上

○目標達成に向けた主な課題及び今後の方向性

- ・機関リポジトリにおける研究論文以外の研究データの登載や、研究データの管理・利活用の方針・計画の策定が進んでいないなど、取組が不十分であり、研究者のデータ管理・利活用の意識や基本的な考え方についての認識も低い
- ・内閣府(科技)は、国際認証基準等に基づくリポジトリの整備・運用のガイドライン及び国研におけるデータポリシーの策定を促進するためのガイドラインを2018年度に策定
- ・研究データの特性等を踏まえて研究データを保存・公開するためのリポジトリの整備や研究データの管理・利活用のための方針・計画の策定を促進し、データインフラを通じた機械判読可能性と相互運用性の確保、諸外国の研究データ基盤との連携を促進
- ・研究者や大学・国研等における現状・取組等についての調査・分析を行い、研究者等の意識向上等に資する方策を検討

³⁵ データの特性から公開すべきもの(オープン)と保護するもの(クローズ)を分別して公開する戦略。

³⁶ 研究成果(論文等)の根拠となるものを含む。

³⁷ データインフラのうち、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。

³⁸ リポジトリのうち、大学・国研等が管理するもの。我が国では主に大学が管理している。

³⁹ 研究データの管理・利活用のための方針については国研が2020年度末までに策定、計画については競争的研究費による研究実施者が策定することを要請する制度を2021年度予算における公募までに、各府省・研究資金配分機関において導入。

① イノベーションにおけるオープンサイエンスのためのデータ基盤の必要性・重要性

活版印刷技術の登場により、本や学術ジャーナルの大量印刷が可能となって以降、物流や知的財産制度の整備とともに、よりオープンな知的基盤が構築された。その結果、多くのイノベーションが創出され、人類の発展をもたらした。

これまで大量印刷と物流による知的基盤が支えてきた科学は、ICTの発展により、サイバー空間が支える科学へと大きく変容し、学界、産業界、市民等あらゆる者がサイバー空間にある研究データを利活用し、協働によって知の創出をするというオープンサイエンスが進展してきている。このような社会の変化に応じて、新たな制度を整備しつつ、研究データの取扱いについての対応方針や運用を再定義することが求められる。

したがって、今後も我が国の研究や産業をますます発展させるべくイノベーションを創出するためには、社会インフラとして、オープンサイエンスのためのデータ基盤の構築が必要である。

② 現状認識

現在、ICTを活用した研究データの利活用による科学研究の変容と新しいイノベーション基盤づくりの議論が国際的に進展している。

例えば、米国では、研究資金配分を行う 22 の連邦政府機関全てがそれぞれの所管分野や所管する政府研究機関における研究成果の利活用のための計画を作成済みである。また、EUでは、既存の研究データの流通基盤の統合を進めている。さらに、国際コンソーシアムにおいて、研究者・技術者等による研究データ流通に係る国際標準等についての議論も進んでいる。

我が国では、機関リポジトリは世界でも最多⁴⁰であるが、研究論文以外の研究データの掲載が進んでいないこともあり、機関リポジトリを活用した研究データを管理・公開・検索するプラットフォームを構築するためのシステム開発⁴¹が進められている。

また、研究データを管理・利活用する上で欠かせないデータポリシー⁴²を策定した国研が 2017 年までに 2 法人⁴³にとどまるなど、研究分野別・組織別の特性を考慮したデータポリシーの策定が遅れており、研究データの管理・利活用が進んでいない。

さらに、データマネジメントプラン⁴⁴の策定の要請など、競争的研究費による研究において研究実施者に適切な研究データの管理・利活用を促す仕組みの導入が十分ではなく、研究者のデータ管理・利活用の意識や基本的な考え方についての認識も低い状況にある。

このままオープン・アンド・クローズ戦略を検討せずに研究データの公開が進み、我が国の産業振興等のために優先的に研究データを利活用できる機会を失い、他国の企業等が先んじて商業化等に利活用することがないように、データポリシーやデータマネジメントプランの策定を急ぐ必要がある。

⁴⁰ 各国の機関リポジトリ数は日本 744、米国 500、イギリス 255、ドイツ 244 と続く（N I I 調べ）。

⁴¹ 開発主体はN I I。

⁴² 研究データの管理・利活用についての方針を組織として取りまとめたもの。

⁴³ J A M S T E C、N I E S の 2 法人。

⁴⁴ 競争的研究費による研究実施者が策定する研究データの管理・利活用についての計画。

③ 今後の方向性及び具体的に講ずる主要施策

研究データを我が国のデータインフラから公開できるよう、主として機関リポジトリを対象としたシステム開発や、国際認証基準⁴⁵等を参考にしたりリポジトリの整備・運用を進める。

その上で、研究分野の特性等を踏まえたオープン・アンド・クローズ戦略を考慮したデータポリシーやデータマネジメントプランの策定を促進し、これらに基づく研究データの管理・公開等を促進するとともに、公的資金による研究成果としての研究データについては、データインフラを通して機械判読可能化を促進する。

さらに、これらの取組が大学・国研等で適切に行われるよう、研究データの管理や公開・共有に従事する研究者等の意識向上や基礎的な知識の習得のための取組や、研究者や大学・国研等における現状、取組等についての調査・分析を行い、研究者等の意識向上等に資する方策を検討する。

i) リポジトリの整備及び展開

【科技、文】

- ・文部科学省が主体となり、機関リポジトリを活用しクラウド上で共同利用できる研究データの管理・公開・検索を促進するシステムを開発し、2020年度に運用開始
- ・文部科学省が主体となり、全文データベース⁴⁶に登載された論文と識別子⁴⁷を付与した研究データを紐づけ、管理・公開するシステムの開発を2018年度中に検討
- ・国際認証基準等に基づくリポジトリの整備・運用のガイドライン（公開データの検索可能化、諸外国の研究データ基盤との相互運用性等を含む。）を内閣府（科技）が策定し、大学・国研等にガイドラインの適用を推奨
- ・ドメイン語彙策定への関係機関の参画等を通じて分野間データ連携基盤と連携

ii) 研究データの管理・利活用についての方針・計画の策定等

【内閣官房、科技、食品、総、文、厚、農、経、国、環、防】

- ・内閣府（科技）は、国研におけるデータポリシーの策定を促進するためのガイドラインを2018年6月までに策定
- ・国研は、研究分野の特性、国際的環境、産業育成等に配慮し、必要に応じてオープン・アンド・クローズ戦略を取り入れ、データポリシーを策定⁴⁸
- ・競争的研究費制度の目的、対象等を踏まえ、大学・国研・企業等の研究実施者がデータマネジメントプラン等のデータ管理を適切に行う仕組み⁴⁹を、各府省・研究資金配分機関が所管の競争的研究費制度に導入⁵⁰（ガイドライン策定や公募要領改訂等）
- ・データポリシー・データマネジメントプランに基づく公的資金による研究データの管理・公開等を促進し、公的資金による研究成果としての研究データについては、

⁴⁵ ここでは、2017年9月に設立されたりリポジトリの国際的な認証機関であるCTS（Core Trust Seal）が定めた認証基準を指す。

⁴⁶ 論文等の書誌（タイトル、著者名、発行年等）のみでなく、全文を登載し、閲覧のため提供する情報の集まり。

⁴⁷ 研究成果に対し、それを一義的に識別し、国際的に通用するデジタル情報。

⁴⁸ 策定法人は、2017年末時点で2法人であるが、2020年度末までに24法人（国研のうち、研究資金配分機関であるAMED、JST、NEDOを除く。）全てでの策定を目指す（法人内の特定の研究センター・部門等の範囲で策定した場合を含む。）。

⁴⁹ データマネジメントプランの策定、指定したりリポジトリでの公開、データ作成者名等の報告書等での明示等。

⁵⁰ 導入府省・研究資金配分機関は、2018年度当初時点で4であるが、2021年度予算における公募までに14府省・機関全てでの導入を目指す（制度内で特定のプログラム・事業等で導入した場合を含む。）。

データインフラを通して機械判読可能化を促進

iii) 人材の育成及び研究データ利活用の実態把握

【科技、文】

- ・2017年度に開発・公開された基礎的な能力開発プログラム⁵¹について、専門性を高めた拡張版の開発・公開及び関係機関における受講の促進
- ・大学・国研等のデータポリシー等作成状況、リポジトリ整備状況、研究データ利活用の優良事例、研究者の公的資金による研究データの公開状況や利活用上の障壁・促進要因等の調査を2019年度から実施⁵²

⁵¹ 主に大学の研究支援職員が研究データの管理に関する基礎的な知識の習得を目的とし、オープンアクセスリポジトリ推進協会が開発したもの。

⁵² N I S T E P「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[調査資料-268]」(2017年12月)では、科学技術の専門家(回答者は約1400名)を対象として、先行的な調査を実施している。