

オープンサイエンスの推進について

文部科学省

研究振興局参事官(情報担当)付

科学技術・学術政策局科学技術・学術戦略官(制度改革・調査担当)付



文部科学省

MEXT

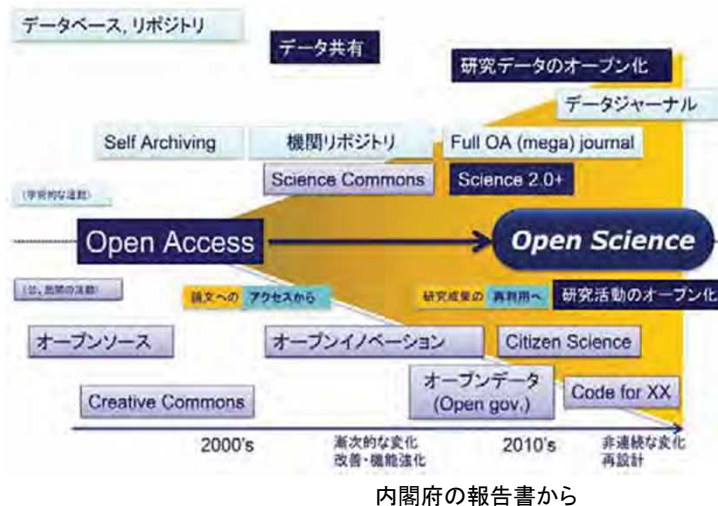
MINISTRY OF EDUCATION,
CULTURE, SPORTS,
SCIENCE AND TECHNOLOGY-JAPAN

I. オープンアクセスからオープンサイエンスへ

「科学技術基本計画」の記載が変化し、「オープンサイエンス」が大きなトピック

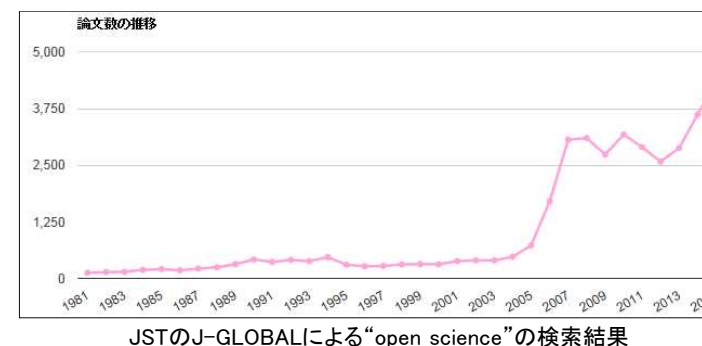
第4期(H23):「研究情報基盤の整備」

「国として、研究成果の情報発信と流通体制の一層の充実に向けて、研究情報基盤の強化に向けた取組を推進する。」



第5期(H28):「オープンサイエンスの推進」

「国は、資金配分機関、大学等の研究機関、研究者等の関係者と連携し、オープンサイエンスの推進体制を構築する。公的資金による研究成果については、その利活用を可能な限り拡大することを、我が国のオープンサイエンス推進の基本姿勢とする。」



「論文」の公開(オープンアクセス)

研究資金配分機関(我が国はJSPS、JST)は、研究者に、研究成果としての論文を無償公開するよう推奨。→商業出版誌で公表した論文は、出版社の許諾を得た上で、所属機関のリポジトリ(大学図書館が運営する論文公開サイト)等で無償公開。

リポジトリに掲載されたデータはH19:30万件→H26:200万件に増加。

今後、公開を更に徹底する方向で調整を進めている。

オープンアクセスから「研究データ」を含めたものとして拡大

左に加え、研究資金配分機関が、研究者に、データの登録・公開を進めている。

・科学技術振興機構: データマネジメントプランに基づき研究データの保存・公開を実施。

・日本医療研究開発機構: 研究データの登録先を指定し、公開
また、海洋研究開発機構等の研究開発法人も、「データポリシー」を定め、データ公開を実施。

今後、公開を更に拡大することが課題

II. オープンサイエンスをめぐる国際的動き

(政府レベル)政策レベルの議論が進展

○G8/G7

- ・ G8科学技術大臣会合(英国):各国で研究成果のオープンアクセスを拡大させる方針が確認(2013)
- ・ G7科学技術大臣会合(日本):研究分野の特性にも配慮した上で、オープンサイエンスの推進を決意。日本とEUを共同事務局とする作業部会の設置(2016)

○米国

- ・ NIH:査読論文の指定リポジトリへの掲載を義務化(2008)
- ・ 科学技術政策局の指令で、NIHと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推進(2013)

○英国:

- ・ 英国研究会議
 - オープンアクセスポリシーを公開(2006/2012 改訂)
 - 「データポリシーに関する共通原則」を公表(2011/2015 改訂)

○ドイツ

- ・ ドイツ研究振興協会:オープンアクセスジャーナル投稿料を助成することでオープンアクセスを推進(2010-)

○EU

- ・ 欧州委員会
 - Horizon2020(EUの研究開発・イノベーションプログラム)におけるオープンアクセスガイドラインを制定(2012)
 - 欧州オープンサイエンスクラウド計画を公表(2015)

(関連する国際機関) 頻繁に標準化・規格化が議論

○国際科学会議の委員会(ICSU-WDS World Data System):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(2008-)

○国際コンソーシアムRDA(Research Data Alliance)

- ・ データ共有の持続性確保、信頼性確保、システム化などの国際標準を議論(2013-)
- ・ JSTが日本での総会開催を主催し、データ共有の議論を喚起(2016)

○OECD

- ・ 公的資金によって得られたデジタルな研究データへのアクセスを推進するための原則とガイドライン発表(2007)
- ・ GSF(Global Science Forum)では、オープンサイエンスに関する検討グループが発足(2015)

(海外出版社)

○海外大手出版社は、論文に付随するデータを、海外の指定するリポジトリに登録し、公開することを推奨

Ⅲ.「学術情報のオープン化の推進について」(学術情報委員会審議まとめ 平成28年2月)

公的研究資金による研究成果を対象とし、大学等及び学協会の活動とこれらを支援する関係機関(研究資金配分機関、JST、NII)の取組を検討範囲として審議。

(基本的考え方)

1. 研究成果の公開を通じた利活用を促進→分野を越えた新たな知見の創出

自然科学・人文科学・社会科学を通じて、新たな知見の創出や効率的な研究の推進等に資する。また、研究成果への理解促進や研究成果の更なる普及につながる。

2. 研究の透明性の確保、研究費の効率的な活用(研究の過度な重複を避ける)

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で研究費を効率的に活用できる。

3. 公的研究資金による研究成果を広く社会に還元

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、そのオープン化の必要性は一層強い。

→ これらの意義や国際的動向を踏まえ、文科省として「**公的研究資金による研究成果のうち、論文とそのエビデンスとしての研究データは、原則公開とすべき**」と方針を明確化。

同時に、研究資金配分機関、JST、NII、大学、学協会等において取り組むべき事項について提起(主な取組は次頁参照)。

III.「学術情報のオープン化の推進について」(続き)

(1)論文のオープンアクセスについての取組

○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中に実行。

(2)論文のエビデンスとしての研究データの公開

- 公的研究資金による研究の実施段階から終了後に至るまで、研究データが利活用可能な状態で適切に管理される必要。
- 研究資金配分機関は、申請者に、必要に応じデータ管理計画の提出を求める。また、データの公開について推奨。
- 研究データの公開は、既に分野別の公的なデータベース等がある場合、これらへの掲載を促進。
- 大学等は、研究データの保管に係る基盤整備について、アカデミッククラウドを構築し、活用。
- 公開すべきエビデンスとしての研究データの範囲とその様式について、学協会等は、研究上の必要性等を考慮した検討を行う。
- 日本学術会議は、研究者コミュニティとしてのコンセンサスを形成。

(参考)

- 日本学術会議:「オープンサイエンスのあり方に関する提言」を公表(平成28年7月)
- ー 研究分野を越えた研究データの管理およびオープン化を可能とする研究データ基盤の整備
- ー 研究コミュニティでのデータ戦略の確立
- ー データ生産者およびデータ流通者のキャリア設計

(3)研究成果の散逸等の防止

- 大学等は、研究成果の管理に係る規則を定め、散逸、消滅、損壊を防止。
- 研究成果にデジタル識別子(DOI)を付与し管理する仕組みを確立。

(4)研究成果の利活用

- 学協会は、学術誌に掲載された論文に係る著作権ポリシーや研究データの利用ルールを明示。
- 大学等や研究者コミュニティにおいては、研究データの被引用を、データ作成者の業績として評価することを実行。

(5)人材の育成及び確保

- 大学等において、研究データを専門的に取り扱える人材の育成と確保が必要。
- 関連学協会:日本学術会議によるアンケートでは、
 - ー 過半の学協会が、論文・データ、データベースを既に公開
 - ー 半数程度の学協会が、データ項目の測定条件を共通化し、一層価値が高まる可能性を持つデータもあると回答
 - ー しかし、オープンサイエンスに関するシンポジウム・ワークショップなどの開催実績がある学協会は1割程度であり、まだ十分に浸透しているとは言えない。

IV. これまでの審議等を踏まえた施策の実施について

施策の実施に当たっては、公的研究資金による研究成果（論文及び研究データ）の利活用を拡大することを基本とする。

このため、国際的な動きや国内におけるオープン化をめぐる状況を踏まえ、具体的な施策を進める上での方向性や留意すべき点等について、御審議いただきたい。

（主な観点）

1. 海外では研究資金配分機関による公開義務化が進展している。我が国の競争的研究費における、研究データの公開・共有への対応についてどのように進めるべきか。
2. 一部の分野や研究プロジェクトにおいては、データ公開・共有が進んでいるが、このような取組を、他分野や機関の取組として拡大していく必要がある。このような状況を踏まえ、大学・研究機関におけるデータ基盤の整備、データの管理及び公開の取組をどのように進めるべきか。
3. 研究データのオープン化においては、データ公開を促進する一方で、機密保持等の観点から公開になじまないデータの取扱いを規定する必要がある。各分野ごとのデータの公開・非公開の考え方や公開のあり方等について、どのように研究コミュニティの議論を喚起し、実行するか。
4. 研究データの公開・共有を進め、その利活用を推進していくに当たっては、データの利活用等に精通した人材が重要な役割を果たす。専門的な技能を持つ情報人材の育成・確保をどのように進めるべきか。

V. 具体的施策について

1) 競争的研究費への導入検討

◆ 現状・課題

AMEDやJSTにおける一部の競争的研究費においては、当該プログラムとしてのデータ管理に関する方針を策定するとともに、研究チーム毎にデータ管理に関する計画を策定し、研究により得られたデータの共有・公開を促進しているが、他の大部分の競争的研究費では特段の対応はとられていない。

◆ 今後の方向性

先行事例を参考にしつつ、文部科学省及び関係法人の他の競争的研究費においても、データの共有・公開への対応について検討を行っていく。検討に当たっては、公開を前提としつつ、データの特性や内容に応じて、公開／非公開の別について柔軟に対応できるような仕組みとすることが重要である。

文部科学省としては、第5期科学技術基本計画期間中に導入が進展するよう取組を進めるとともに、対応状況のフォローアップ等を行っていく。

(参考 1) 競争的研究費への導入事例

(1) AMEDにおける事例



- ◆ AMEDの「疾病克服に向けたゲノム医療実現化プロジェクト」では研究者や研究参加者等に関係する権利保護と、ゲノム情報の共有による関連分野の研究の推進を両立させるための**データシェアリングポリシー**を策定。共有・公開の範囲を「制限共有」「制限公開」「非制限公開」の3つに分類。
- ◆ 当該プロジェクトが資金提供を行う研究課題にて発生するゲノムデータ及び臨床情報や解析・解釈結果等を含めたゲノム情報が対象。データの種類・規模、登録データベース、登録時期、公開・共有の範囲等を記載した「**データマネジメントプラン**」を公募時に提出。**公募審査の評価対象**となる。
- ◆ データは、原則として、**ゲノム解析終了後2年又は論文採択時のいずれか早い時点**で、AMED指定の公的データベースに登録。

(2) JSTにおける事例



- ◆ 「**戦略的創造研究推進事業におけるデータマネジメント実施方針**」に基づき、平成 28 年度以降に新たに設定されたCREST・さがけ研究領域等で採択された研究代表者は、**研究チームの成果として生じる研究データの保存・管理、公開・非公開の区別、及び公開可能な研究データの運用指針**をデータマネジメントプランとして定める。
- ◆ 研究データが個人情報保護・機密保持・商業化・国家安全保障の観点から共有されるべきものではない、もしくは共同研究契約等で成果の公開に制限がある等、考慮すべき事情があれば研究データを**非公開**とすることは妨げない。

2-1)研究データの保管に係る基盤整備

- ◆ 現状・課題

一部の分野においては国立研究開発法人や大学等においてデータリポジトリ等の整備が進められているが、さらに他機関も含めてその充実を図る必要がある。
- ◆ 今後の方向性

データを集積するリポジトリ等の基盤については、DIAS等特定の研究プロジェクトや分野の中核機関である国立研究開発法人等において既に整備されているものがあれば、それらの活用を促進していく。一方、これ以外の分野の研究データについては、研究者の所属機関におけるデータリポジトリを活用することを基本とする。

このため、国立研究開発法人において、それぞれの研究分野のデータプラットフォームの機能を備えるべく基盤の整備を図る。また、各大学等において、NIIと連携のうえアカデミッククラウドを構築、活用する。

(参考2) データリポジトリ等の基盤整備等に関する施策例

施策名	29年度 概算要求額(億円)	28年度 予算額(億円)	概要	計上 主体
データプラットフォーム 拠点形成事	57(新規)	—	特定国立研究開発法人をはじめとした国立研究開発法人において、我が国が強みを活かせるナノテク・材料、ライフサイエンス、防災分野で、膨大・高品質な研究データを利活用しやすい形で集積し、産学官で共有・解析することで、新たな価値の創出につなげるデータプラットフォーム拠点を構築。	MEXT
研究成果を利用する ための基盤整備	JST運営費交付金 の内数(新規)	—	データ駆動型科学の実施に向け、JST戦略的創造研究推進事業と連携し、研究データ管理計画の提出を求める領域、プログラムを対象とした、データシェアリング用のリポジトリを構築し、データの共有・公開を進める。	JST
ライフサイエンスデー タベース統合推進事 業	JST運営費交付金 の内数	JST運営費交付金 の内数	我が国におけるライフサイエンス研究の成果が、広く研究者コミュニティに共有され、活用されることにより、基礎研究や産業応用につながる研究開発を含むライフサイエンス研究全体の活性化に貢献するため、国が示す方針の下、様々な研究機関等によって作成されたライフサイエンス分野データベースの統合に向けた、戦略の立案、ポータルサイトの構築・運用及び研究開発を実施し、ライフサイエンス分野データベースの統合を推進する。	JST
オープンサイエンス研 究データ基盤の整備	国立大学法人等運 営費交付金の内数 (新規)	—	国立情報学研究所が大学等と連携して、大学等のデータ基盤に関する主要なシステム開発を実施する。	NII

2-2) 研究データの散逸等の防止や利活用について

◆ 現状・課題

大学等で研究データの管理に係る規則を定め、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とともに、データの利活用を促進するために、デジタルオブジェクト識別子(DOI※)を付与する仕組みを構築する必要がある。

◆ 今後の方向性

研究成果の散逸等を防止するために、大学や国立研究開発法人等においては、データ管理に係る規則等の整備を推進する。また、研究データにDOIを付与する仕組みの構築を推進する。

文部科学省としても、第5期科学技術基本計画期間中に関係機関におけるルールの整備が行われるよう、整備促進に向けた状況調査や盛り込むべき内容の提示等を行っていく。

※ DOI: データに付与される国際的な識別子。公開されたデータへの永続的なアクセスを保障するために必要となる。

(参考3) DOI付与に関するJaLCの取組の事例

(1) ジャパンリンクセンター (JaLC) における事例

- ◆ JaLCは日本で唯一のデジタルオブジェクト識別子 (DOI) 登録機関であり、日本の情報提供機関および研究機関 (NDL, NII, JST, NIMS) による共同運営。
- ◆ 研究データへのDOI登録に特有の課題抽出とその解決、運用方法の確立、DOIの活用方法などの検討を行っている。
- ◆ また、研究データへのDOI登録実験プロジェクトで得られた知見と議論をもとに「研究データへのDOI登録ガイドライン」として取りまとめた。今後、研究機関等がデータに対するDOI登録を開始する際の指針になり、実作業への参考となることを目指している。

2-3) データ共有/公開に対する評価の取組について

◆ 現状・課題

研究データの共有・公開の推進は研究者をはじめとする関係者の理解を得て進めることが基本となるため、研究者コミュニティとの連携の下、オープン化の推進に向けた研究者のモチベーションやインセンティブを高める方策について取り組む必要がある。

◆ 今後の方向性

データの利活用を促進するために、データの公開・共有に係る取組や有用なデータの産出について、各機関や研究者の業績評価に適切に反映されるよう各機関の評価指針等において明確化していく。

また、学協会や研究データ利活用協議会等の団体とも連携し、オープンサイエンスの推進に係る情報発信の取組を行っていく。

(参考4) 関連する取組の事例

(1) 「研究データ利活用協議会」での検討

- 研究データの利活用に係る議論を促進するため2016年6月に設立。JST、NIMS、NII、NDL、NICT、AIST等をはじめとする研究機関等や個人が参加。ジャパンリンクセンター(JaLC)が事務局を務める。
- 年3~4回程度研究会、年1回程度報告会を実施するなどして、オープンデータのマインド醸成や国の政策への反映を目指す。

(2) JST科学技術情報委員会において「わが国におけるデータシェアリングのあり方に関する提言」を取りまとめ



【抜粋】提言4: 研究データに係る研究業績評価の仕組みを見直すべき

- 研究者や研究組織のデータ共有意識を高揚させるため、従来の論文発表による評価と同様に、研究データ作成と提供を成果として評価対象とし、研究データシェアリングが研究プログラムや研究組織の業績として評価されるよう、内閣府及び各省の研究開発評価指針を改定すべき
- 改定にあたっては、データ研究、データ基盤構築に係る研究業績の新たな評価軸、研究費の一定額をデータ基盤整備等に投入できる枠組みの導入などが必要

3) 分野の特性に応じたデータの公開／非公開のあり方の検討

◆ 現状・課題

データの共有・公開の際の公開/非公開の区分については、多くの機関や研究プロジェクトにおいて、一般公開(オープン)と非公開(クローズ)の間に、制限共有(セミクローズ)/制限公開(セミオープン)のいずれかあるいは両方を設けている。また、公開までの猶予期間を設けているケースもある。

◆ 今後の方向性

データの共有・公開に当たっては、オープンとクローズの間の中間的な取扱いを可能とし、具体的な取扱いについては、データの特性や内容、研究機関の方針等を踏まえ、研究プロジェクトにおけるデータ管理計画や研究機関におけるデータマネジメントポリシー等において定めていく。特に、公開等になじまない性格のデータ(機密保持、企業秘密、国家安全保障に関わるデータや商用化・産業化を目的としたデータ、共同研究契約などで研究成果の公開に制限があるデータなど)については、非公開とするかまたは特別な配慮のうえで公開等を制限する。また、公開までの猶予期間(エンバーゴ)を適切に規定する。

文部科学省としては、第5期科学技術基本計画期間中に関係機関・研究プログラム等において整備が進展するよう取組を進めるとともに、各機関等における対応状況のフォローアップ等を行っていく。

(参考5)大学・研究機関等におけるデータ共有の事例について

※10月18日 第8回学術情報委員会資料を一部更新

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)
大学、大学共同利用機関法人、国立研究開発法人による取組例			
筑波大学計算科学研究センター	全国の計算素粒子物理研究の成果であるQCD真空配位(格子QCDシミュレーションの基礎データ) 【データベース】 JLDG QCD ensemble	①セミオープン(制限公開) データグリッド JLDG、及び、JLDGと連携する国際データグリッド ILDG のユーザに限定して公開(当該分野の研究者・学生は、原則、申請によって、ユーザーアカウントを取得できる)。データを利用した研究の成果を公表する際は、データを提供した研究グループが指定する論文を引用すること、としている。	研究グループがデータを公開するか否か、公開する場合の公開時期は、研究グループの判断にゆだねている。当該データを用いた研究の最初の本論文が発表された時点で公開する事を推奨している。

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)
大学、大学共同利用機関法人、国立研究開発法人による取組例			
京都大学人文科学研究所	漢代から中華民国初期にわたる約5,000枚近い拓本の全文画像及び全ての文字画像 【データベース】 拓本文字データベース	①オープン(公開)	データベースに登録したものから即時公開。
大阪大学蛋白質研究所	生体高分子の立体構造に関するデータ 【データベース】 Protein Data Bank Japan (PDBj)	①オープン(公開)	データバンク方式であり、データベースへの登録者が登録時に下記の中から公開時期を選択する。 1 データベースに登録したら即時公開。 2 論文発表(bioRxiv等のPublic Preprintを含む)があったら公開(最長1年) 3 期日を指定(登録から4週間後、6週間後、8週間後、6か月後、1年後)して公開。
人間文化研究機構国文学研究資料館	国文学研究資料館が調査研究等の活動により作成する古典籍のデータ及び国内外の大学等研究機関が所蔵する日本文学を含む広範な分野・領域の古典籍のデータ 【データベース】 日本古典籍総合目録データベース 館蔵和古書目録データベース	①オープン(公開) ただし、国文研以外の機関が所蔵する古典籍のデータについては、転載などの利用に関しては各所蔵者の許諾が必要。また、国文研所蔵古典籍データの一部は国立情報学研究所情報学研究データリポジトリからもCC BY-SAの条件で公開。	公開用データベースに登録したものから即時公開。
自然科学研究機構国立天文台	国立天文台及び大学の観測装置で共同利用等により取得した天文観測データ 【データベース】 SMOKA すばる望遠鏡アーカイブシステム(STARS) など	①オープン(公開) ②セミオープン(制限公開) SMOKAについて、観測データを取得する場合はユーザ登録が必要。 ③セミクローズ(制限共有) STARSについては、共同利用者に対して観測データを共有。 の3分類	観測装置によって、観測者の占有期間1～2年間を設定し、その後、当該観測データを公開。
情報・システム研究機構国立遺伝学研究所	DNA塩基配列と機能注釈データ 【データベース】 DNA Data Bank of Japan (DDBJ)	①オープン(公開)	登録者は公開予定日を設定し延長することができる。

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)
海洋研究開発機構 (JAMSTEC)	調査観測データ、シミュレーションデータ、画像、 図面等の機構が保有する研究設備、施設を使った 調査研究で得られた各種情報及びそれらを記録し たもの 【データベース】 深海映像・画像アーカイブス (J-EDI) 地震研究情報データ提供システム (J-SEIS) など	①オープン (公開) 公開猶予期間後に公開。調査実施時の各 種条件・取決め・契約等がある場合は個 別対応。 ②クローズ (非公開) 個別に判断 (事故調査に協力した映像や 画像などのデータ) の2分類	取得した研究者が優先的 に利用できる期間 (公開 猶予期間: 原則2年) を設 定し、経過後に公開。 航海に関する情報や航海 報告等は2ヶ月で、定常的 に取得する観測項目につ いては準備でき次第公開。 なお、調査実施時の各種 条件・取決め・契約等が ある場合は個別対応。
物質・材料研究 機構 (NIMS)	無機材料や高分子材料、金属構造材料の物質構造 等、NIMSの研究活動で得られたデータに加え、 論文等からのデータを多数含む。 【データベース】 物質・材料データベース (MatNavi)	①オープン (公開) NIMSにおけるサイトポリシーに同意してい るものとして原則公開としている。	データベースに登録した ものから即時公開。
理化学研究所	理化学研究所における研究活動で取得した哺乳類 ゲノムの機能注釈に関するデータや天然化合物の名 称・構造、生物活性情報のほか、物性データや機器 分析データなど 【データベース】 FANTOM NPEDIA など	①オープン (公開) ②クローズ (非公開) NPEDIAでは、独自のスクリーニング結果、 共同研究で得られた評価結果については 非公開。 の2分類	データベースに登録したも のから即時公開。 FANTONについては、プロ ジェクト主論文掲載時にデー タを公開。

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)
研究プロジェクトによる取組例			
情報統合型物質・材料開発イニシアティブ (MI ² I)	NIMSの研究活動で得られた物質・材料の基礎データや、データ科学的手法等の解析ツールを格納。	①セミクローズ(制限共有) プロジェクト期間中は、プロジェクト参加者とコンソーシアム会員の範囲内でのセミクローズ(制限共有)。 ②クローズ(非公開) データ所有者の判断により、データ所有者の独自データ等、共有できないデータ・ツールは、クローズ(非公開)。 の2分類	プロジェクト参加者とコンソーシアム会員の範囲内での制限共有を実施(～2020.3.31まで) 公開については、今後検討。
データ統合・解析システム (DIAS)	地球環境情報のプラットフォームとして、国内外の各機関が保有する衛星観測、大気・地上気象観測、海洋観測、気候変動予測モデルの出力データなど	①オープン(公開) DIAS利用規約に同意することで利用可能 ②セミオープン(制限公開) データ利用者の申請を受け、データ提供者の許可を得ることで利用可能 ③セミクローズ(制限共有) データ提供者があらかじめ設定したグループの構成員のみ利用可能 ④クローズ(非公開) の4分類	データベースに登録したものから即時公開。
バイオサイエンスデータベースセンター (NBDC)	公的資金を用いたプロジェクト等で産生された生命科学分野の研究データ(ゲノム配列、タンパク質、代謝産物、画像等)及び公的資金を用いたプロジェクト等で産生されたヒトに関するデータ (DDBJ、PDBjと連携)	①オープン(公開) ②セミオープン(制限公開) NBDCヒトデータベースで制限公開データを取得する場合には、利用申請が必要。 の2分類	論文等による成果公開や知的財産権取得等の期間の後、公開。

4)人材の育成及び確保

◆ 現状・課題

研究データの共有・公開を進め、その利活用を推進していくに当たっては、データの利活用に精通した人材が重要な役割を果たすが、現状これらの人材の育成支援は体系的に行われていない。

◆ 今後の方向性

「第4次産業革命に向けた人材育成総合イニシアチブ」に沿って、情報スキルを備えた人材の育成を進める。
また、データの共有・公開に当たっては、データの加工・確認等を行う専門人材やデータベースの開発・整備・運用・品質管理等を行う専門人材も必要である。このため、これらの専門人材を研究プロジェクトや研究機関等において確保するための措置について検討を進めていく必要がある。

(参考 6) 「第4次産業革命に向けた人材育成総合イニシアチブ」関連施策のうち、人材の育成・確保に焦点を当てている施策の例

施策名	29年度 概算要求額 (億円)	28年度 予算額 (億円)	概要	計上 主体
データ関連人材育成プログラム	3(新規)	—	企業等がコンソーシアムを形成し、ポスドクター等(博士過程学生含む)に対する研修プログラムを開発・実施することで、我が国社会で求められるデータ関連人材を育成し、ポスドクター等の高度な研究能力を有する人材の多様な場での活躍を促進。	MEXT
成長分野を支える情報技術人材の育成拠点の形成(enPiT)	22	6.5	大学における情報技術人材の育成機能を強化するため、産学協同の実践教育ネットワークの形成により、学部学生を対象とする課題解決型等の実践的な教育の一層の充実を図るとともに、新たに社会人学び直しのための体系的教育プログラムの開発を推進。	MEXT
大学等における数理・データサイエンス教育の強化	国立大学運営費交付金の内数(新規)	—	世界に先駆けた「超スマート社会」の実現(Society5.0)に向けて、専門分野の枠を超えた全学的な数理・データサイエンス教育機能を有するセンターを整備し、大学における数理・データサイエンスの教育強化を図ることで、数理的思考やデータ分析・活用能力を持ち、社会における様々な問題の解決・新しい課題の発見及びデータから価値を生み出すことができる人材育成を促進。	国立大学法人

（参考 7）「第4次産業革命に向けた人材育成総合イニシアチブ」 ～未来社会を創造するAI/IoT/ビッグデータ等を牽引する人材育成総合プログラム～

- 「第5期科学技術基本計画（平成28年1月閣議決定）」において謳われている「超スマート社会」の実現、及び「理工系人材育成に関する産学官円卓会議における行動計画」等を踏まえ、関連施策の一体的な推進が求められている
- 生産性革命や第4次産業革命による成長の実現に向けて、**情報活用能力を備えた創造性に富んだ人材の育成が急務**
- 日本が第4次産業革命を勝ち抜き、未来社会を創造するために、特に喫緊の課題であるAI、IoT、ビッグデータ、セキュリティ及びその基盤となるデータサイエンス等の人材育成・確保に資する施策を、初中教育、高等教育から研究者レベルでの包括的な人材育成総合プログラムとして体系的に実施**

参考：必要とされるデータサイエンス人材数(※)

- 世界トップレベルの育成（5人/年）
- 業界代表レベルの育成（50人/年）
- 棟梁レベルの育成（500人/年）

- 独り立ちレベルの育成（5千人/年）
- 見習いレベルの育成（5万人/年）

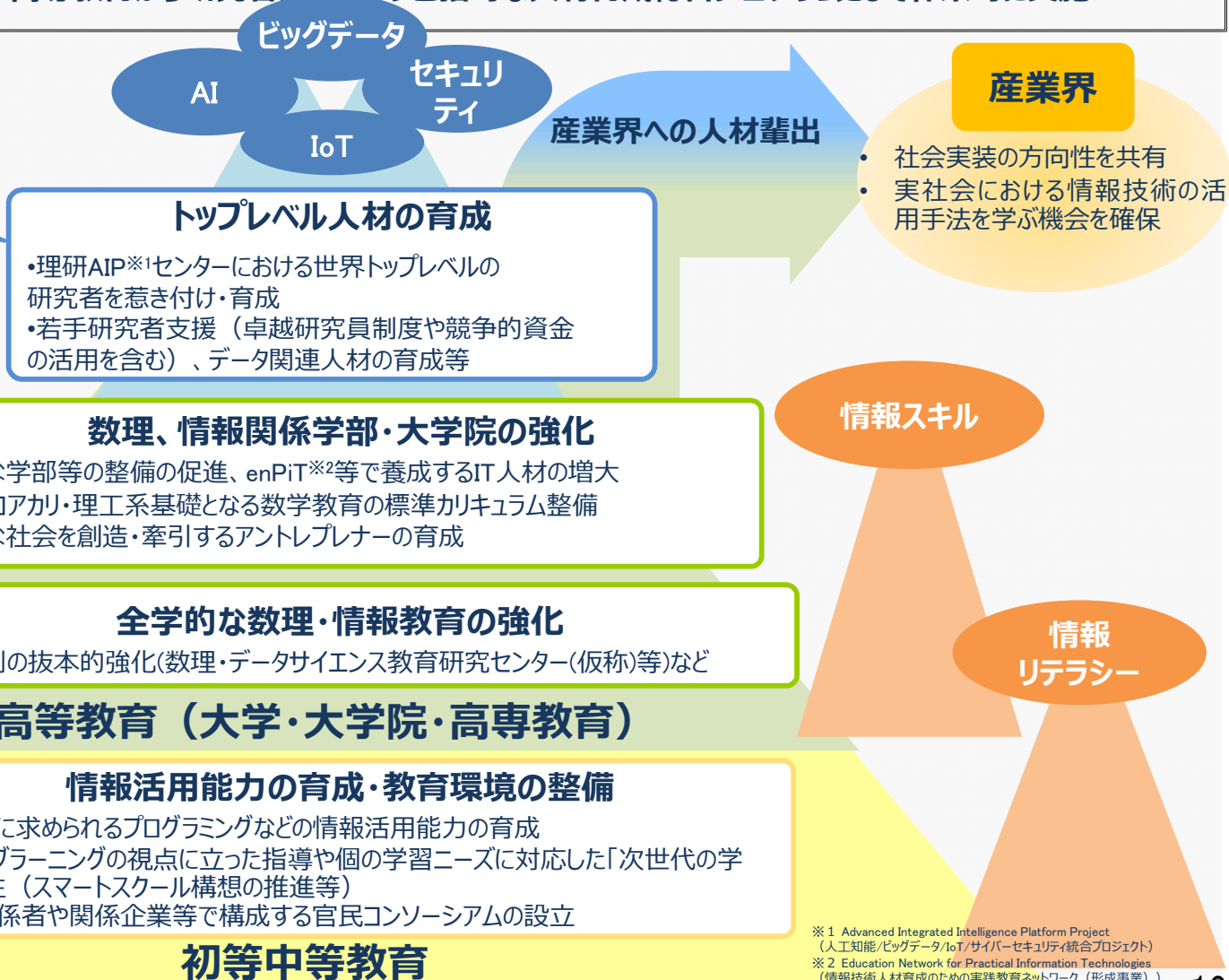
現状（MGIレポート）
日本：3.4千人
US:25千人、中国：17千人

- リテラシーの醸成（50万人/年）

大学入学者/年：約60万人

- 小学校における体験的に学習する機会の確保、中学校におけるコンテンツに関するプログラミング学習、高等学校における情報科の共通必修科目化といった、**発達の段階に即したプログラミング教育の必修化**
- 全ての教科の課題発見・解決等のプロセスにおいて、**各教科の特性に応じてICTを効果的に活用**
- 文科省、経産省、総務省の連携により設立する官民コンソーシアムにおいて、**優れた教育コンテンツの開発・共有等の取組を開始**

高等学校：約337万人（3学年）
中学校：約350万人（3学年）
小学校：約660万人（6学年）



※1 Advanced Integrated Intelligence Platform Project
(人工知能/ビッグデータ/IoT/サイバーセキュリティ統合プロジェクト)
※2 Education Network for Practical Information Technologies
(情報技術人材育成のための実践教育ネットワーク(形成事業))

※注：左吹き出しの人数は「ビッグデータの利活用のための専門人材育成について」（大学共同利用機関法人情報・システム研究機構、平成27年7月）から引用

VI. 参考情報

(参考：データを基盤とした研究プロジェクトの推進について)

◆ 事例1 「地球環境情報プラットフォーム構築推進プログラム」
(平成29年度概算要求額: 19億円、平成28年度予算額: 4億円、計上主体: MEXT)

これまでに開発したデータ統合・解析システム(DIAS)を、企業も含めた国内外の多くのユーザーに長期的・安定的に利用される「気候変動への適応・緩和をはじめとした多様な社会課題の解決に貢献していくための社会基盤」へと発展させるため、安定的運用に必要な設備整備を行うとともに、気候変動適応策・緩和策等に貢献する地球環境情報プラットフォーム活用のための運営体制の整備や共通基盤技術の開発を推進。

◆ 事例2 「情報統合型物質・材料開発イニシアティブ (MI²I : "Materials research by Information Integration" Initiative) 」
(平成29年度概算要求額: 運営費交付金の内数、平成28年度予算額: 運営費交付金の内数、計上主体: NIMS)

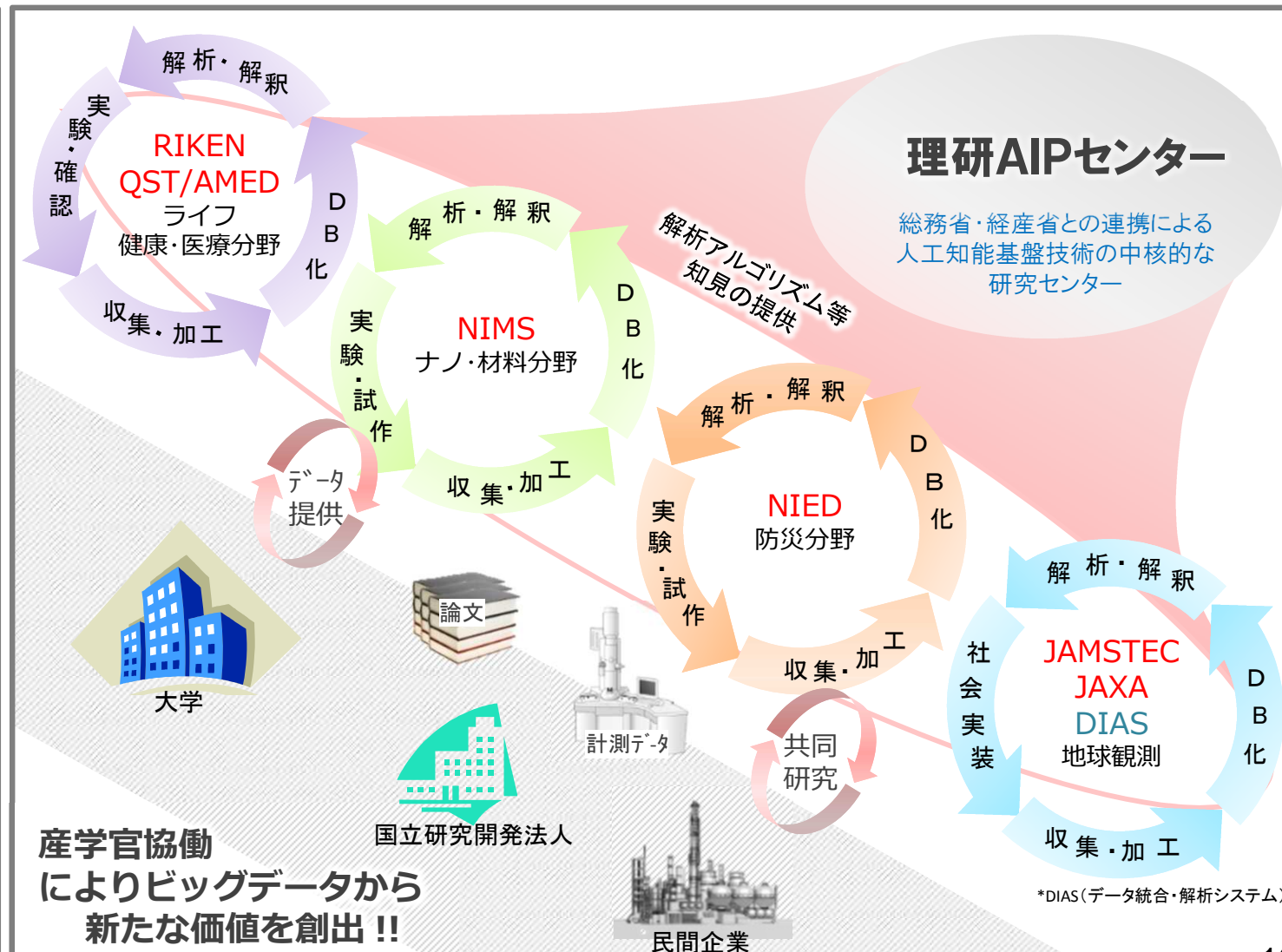
物質・材料研究機構の研究活動で得られた物質・材料の基礎データやデータ科学的手法等の解析ツールを格納。現時点では、プロジェクト参加者とコンソーシアム会員の範囲内での制限共有を実施。

◆ 事例3 国立研究開発法人を中核としたビッグデータ研究拠点の構築

- 将来の成長分野であるAI技術を活用し、我が国が世界の競争に打ち勝つためには、**解析に用いるAI技術の高度化のみならず、解析対象の分野において大量の高品質なデータを整備するとともに、それらを共有・解析するためのプラットフォームの構築**が急務。
- 国立研究開発法人等では様々な分野研究を通じて膨大・高品質なデータを蓄積しており、これらを産学官で活用し、**新たな価値の創造につながる解析を行うことができるプラットフォームを構築**することで、データ解析による新材料設計や生命システムの理解、個人の健康予測、防災技術開発等を強力に推進。

事業概要

- ライフ・健康・医療、ナノ・材料、防災、地球観測の各分野で中核となるデータプラットフォームを5年で構築。
- 各プラットフォームはデータ取得からデータベース化、解析にいたる各段階で理研AIPセンターの革新的技術を活用することにより相互に連携。
- 大学等研究機関及び民間企業等と共同研究契約を結び、双方のデータを利用して新たな価値を創出。
- 共同研究相手の意向に応じてオープン・クローズ・シェアを検討。
- 公開可能なデータベースについては広く利活用に供する。
- 例えば、次世代の材料開発の起爆剤となるアプリケーションを産学官連携のもとで開発し、提供。



(参考:「学術情報のオープン化」に関連する平成29年度概算要求)

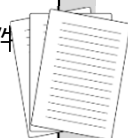


科学技術情報連携・流通促進事業

1. 電子情報発信・流通促進 (総合電子ジャーナルプラットフォーム)

国内の学会の約半数(1,057学会)の計1,948誌の電子ジャーナルを公開するプラットフォーム(4割が英文誌)。

- H11に開始し、270万記事を掲載。
- 年間の論文ダウンロード数
H22: 2,500万件→H27: 7,000万件
(1/3が海外からのダウンロード)



国内外の研究機関・産業界等で幅広く利活用

2. 研究者情報の流通促進 (研究者情報管理)

国内研究者25万人以上の情報を公開するプラットフォーム(大学の研究者32万人のかなりが参加)。

- H10に開始し、研究者情報の国内外への発信に貢献。
- 年間のページビュー
H22: 1,400万件→H27: 2,500万件
(1/4が海外からのアクセス)



3. 基本情報の整備、連携活用システム等の整備 (科学技術総合リンクセンター)

国内資料11,600誌、国外資料3,600誌から書誌情報(論文の基本情報)を整備。

- 現行事業はH21に開始(前身はS30代~)。4,200万件の書誌情報を公表。年130万件の書誌情報を追加。特許情報などの外部データベースとも連携。
- 年間の利用件数
H22: 3,400万件→H27: 8,500万件



科学技術の動向分析や、産学連携等を通じたイノベーション創出の加速に貢献

- これらの持続的整備とともに、「オープンサイエンス」への関心が国際的に高まる中で、それぞれに対応した今日的課題への対応が課題。
(オープンサイエンス: 公的研究資金を用いた研究成果を、容易にアクセスができるようインターネット上で公開し、イノベーションの創出につなげる)

【更なる方策】

- ・ 国際標準の機能を備えることで、我が国の研究成果の発信力を維持・向上。
- ・ 「データジャーナル」の急速な国際動向に対応。

【更なる方策】

- ・ 研究機関や政策立案者のための効果的・効率的な活用に向けた機能強化。

【更なる方策】

- ・ アクセス数が大きく増加しており、これへの対応と、各種データとの関連付けを行う。

- こうした更なる方策を通じて、従来の研究者、学会、産業界による利用に加え、政策立案者、資金配分機関などの幅広い利活用が進展。

【運用システム】

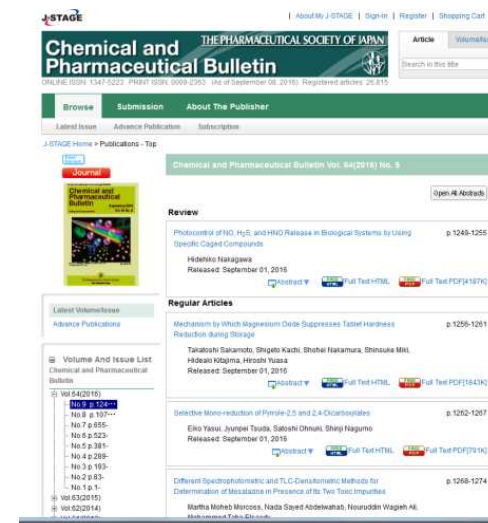
○ 学術論文(電子ジャーナル)を国際的に流通させるためのプラットフォーム。

- ・ 論文を国際的に同定するIDの付与
 - ・ 論文が引用された場合のリンク付与を自動化
 - ・ 学会の投稿審査システム(投稿→査読→採択)をオンライン上で実現
- 国際協調による共通枠組みで実施

○ 各学会における電子ジャーナルの編集・発行・著作権処理の技術サポートを実施。

【現状】

- H11に開始し、270万記事を掲載。国内の学会の約半数(1,057学会)が参加し、計1,948誌の電子ジャーナルを公開(日本循環器学会の”Circulation Journal”をはじめ4割が英文誌)。
- 無償公開を原則。年間の論文ダウンロード数は、H22:2,500万件→H27:7,000万件であり、1/3が海外からのダウンロード。



(ジャーナルトップ画面)

【効果】

- 日本の学会の国際発信力向上に貢献。
「2015年のインパクトファクターが向上した。投稿数も毎月110件以上に増えた」
「J-STAGEに掲載して、海外からの投稿数が飛躍的に増えた」
- 利用者からも高い満足度(93%が「とても役に立った」「役に立った」とアンケートで回答)。
- 海外商業出版社は寡占化が進んでおり、我が国として、それ以外のプラットフォームを通じた電子ジャーナルの流通の仕組みを持つことが重要。

【H29の追加的取組】

- ① 国際標準(又はデファクト機能)を備えることで、我が国の研究成果の発信力を維持・向上。
 - ・ 電子ジャーナルの国際表示規格に対応(国際的な議論を経て2015年に大幅改訂され、各国で対応が進んでいる)。
 - ・ 読者の評価・反応をウェブ上で可視化(国際標準ツールに対応)。
 - ・ 海外データベースとの連携のための著者IDシステムの導入(国際協調が進んでいる)。
 - ・ これらに連動したインタフェースの改良。

② 「データジャーナル」の急速な国際動向に対応。

- ・ 論文のエビデンスとなるデータを「データジャーナル」として公開する動きが急速に進展。我が国として、国際的な議論(※)に参画しながら、エビデンスデータを「データジャーナル」として、国内外で利活用されるよう基盤整備に着手。

データ利用の例:計測技術研究のプロジェクトとして、「化学物質に対する特定の波長を計測したデータ」を取得。このデータは、「リチウムイオン電池材料の長寿命化」という全く異なる研究として活用できる。

※ 既に、国際的な枠組みの中で、データID付与、著者ID付与の標準化が議論されている。今後、リンク付与等、評価の仕組みなども課題となると想定。

2. 研究者情報の流通促進

<http://researchmap.jp/>

researchmap

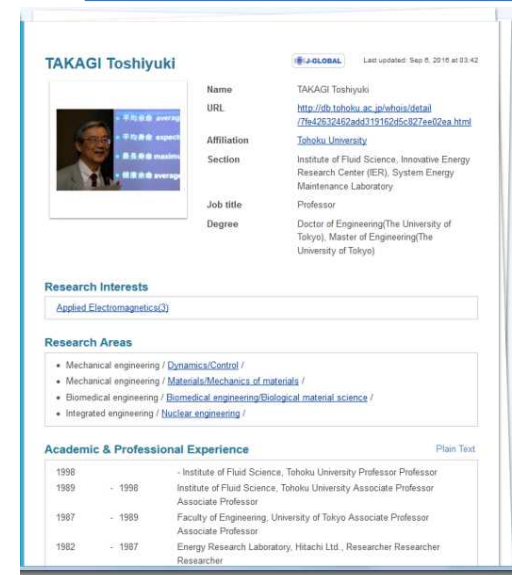
(研究者情報管理:事業名「researchmap」)

【運用システム】

- 研究者の経歴と研究業績(論文、国際会議での発表、受賞歴)等を公開するデータベース。国内研究者の情報を統一的に検索できる。
- 研究者が自身の業績情報を集約できる。データ入力、①研究者が個別入力、②所属機関のデータベースから一括更新、③国際的な論文データベースとの連携による自動更新、で行われる。

【現状】

- H10に開始し、国内研究者25万人以上が掲載(大学の研究者32万人)。我が国の研究者データベースとして最大であり、研究者情報の国内外への発信に貢献。
(年間のページビューは、H22:1,400万件→H27:2,500万件であり、1/4が海外からのアクセス)
- 研究者の事務負担軽減のため、業績リストのデータを(ウィジットという技術を用いて)切り出すことで、競争的資金の申請や業績評価の提出にも容易に活用。



(研究者情報表示画面)

【効果】

- 大学にとって保有データ増とコスト減が実現
「2014年に、研究者データベースのマスターデータをresearchmapとしたことで、論文業績の登録件数が大幅に増加(15万件→34万件)」
「大学の独自データベースを設けるよりも、researchmapをマスターデータとすることで、管理コストが大幅減」
- 研究者が、所属組織を変えても情報が維持できるため、人材の流動化が進展する今日的状況に対応。

【H29の追加的取組】

- 研究機関や政策立案者のための効果的・効率的な活用に向けた機能強化。これにより、例えば、
 - ・「過去5年以内にNatureまたはScienceに論文が採択された女性研究者のリスト」、
 - ・「A大学に所属し、トップクラスの国際会議に論文が採択された特任研究者のリスト」、
 - ・「さがけ研究者が、研究終了後に発表した論文・特許、受賞、新たに獲得した競争的資金のリスト」などの調査ができるようになる。
 - 上記を実現するための技術的対応。
 - ・ 全データ項目を対象とする全文検索の導入
 - ・ データ項目の拡充(大学からのアンケートを踏まえ、学位取得年月、所属学会在籍期間、資格等)
 - ・ 自由入力項目を機械可読とする(コード化)
- (AIによる情報収集と解析機能が別途開発されることを想定しており、その機能と組み合わせることで実現)

3. 基本情報の整備、連携活用システム等の整備

(科学技術総合リンクセンター:事業名「J-GLOBAL」)

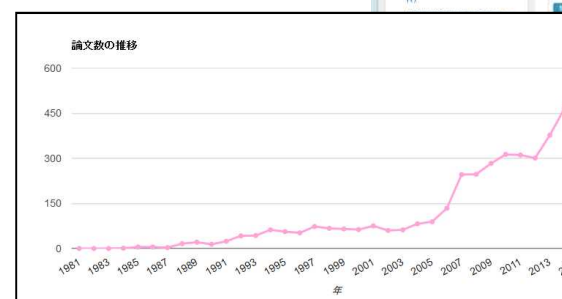
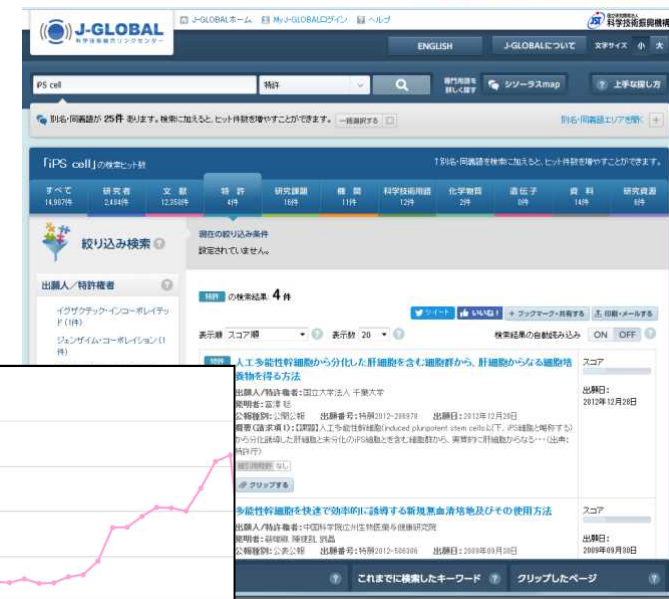
<http://jglobal.jst.go.jp>

【運用システム】

- 国内外の科学技術情報を網羅的に収集・活用するための書誌情報(論文の基本情報)を中心とするデータベース。外部データベース(特許、国会図書館、各種の科学技術情報)との連携で、研究者／文献／特許／研究課題などのデータを一元的に検索。
- 大学・企業が、研究開発や企画立案に先立ち、先行研究の把握や特許情報の収集等に活用。

【現状】

- 現行事業はH21に開始(前身はS30代～)。4,200万件の書誌情報を公表。年間で130万件の書誌情報を追加。
- 国内資料11,600誌(国内誌のほぼ100%)、国外資料3,600誌から書誌情報を作成。分子構造の科学技術データ300万件等も収録。
- 年間の利用件数は、H22:3,400万件→H27:8,500万件に増加。



(検索事例:「機械学習」に関する論文数の推移)

(横断検索結果表示画面)

【効果】

- 大学だけでなく、産業界から高い評価と多くの利用。
- 「研究者や特許などを横断的に検索できるすばらしいサイト。このような画期的なサイトは見たことがない。」
- 「競合他社の調査を行い、製品開発に活用している」
- 「大学との共同研究先を選定する際に活用している」
- 「研究開発には信頼性の高い情報が必須であり、社員にはJ-GLOBALを使うよう指導している」
- 利用者からも高い満足度(89%が「とても役に立った」「役に立った」とアンケートで回答)。

【H29の追加的取組】

- アクセス数が大きく増加しており、その対応と、各種データとの関連付けの対応。
 - ・ 大量の研究データ関連付け処理(データ1件あたり書誌情報との突合処理で4,200万件を実行)のために増強。

オープンサイエンス推進のための研究データ基盤の整備（国立情報学研究所）

【必要性】

- ✓ オープンサイエンスを推進するため、大学で産出される研究データを適切に保存し、利活用できる仕組み（メタデータの標準化、研究データごとに固有の識別子を付ける機能等）を一元的に整備するとともに、公開された研究データから利用者が使いたいデータを簡便に検索できる仕組みが必要。
- ✓ 論文のエビデンスデータが出版社が指定する海外のリポジトリ等に登録されている状況を早急に改善し、研究データが海外のプラットフォームに集積する状況を変えることが必要。
- ✓ 研究公正の観点からも研究データを長期的に保存することが求められており、そのための基盤を効率的に整備することが必要。

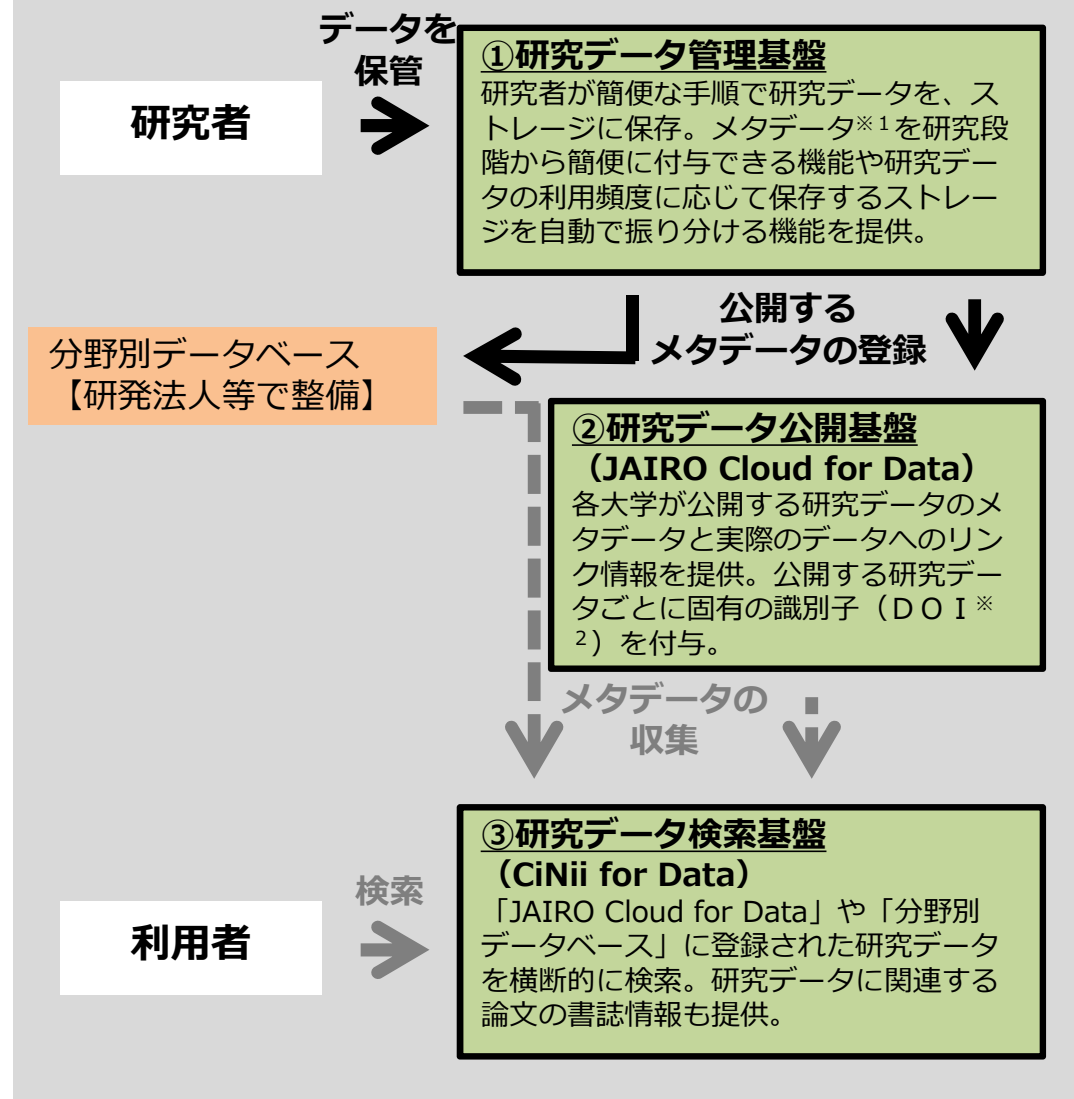
【整備内容】

- ✓ これらの課題に対応するため、クラウド上で共同利用できる研究データ基盤をNIIと大学の連携の下で整備。基盤となるシステム開発はNIIが担い、研究データ保存のために大学が利用するストレージは、共同調達を行うことで効率的に整備。

（システム概要）

- ① 研究者が簡便な手順で研究データをクラウド上のストレージに保存できるようにするシステム【研究データ管理基盤】
 - ② 各大学が保存した研究データのうち、公開するデータに関する情報（メタデータ）と実際のデータへのリンクを提供するシステム【研究データ公開基盤】
 - ③ 研究データ公開基盤や分野別データベースなどに登録された研究データを横断的に検索できるシステム【研究データ検索基盤】
- ✓ データ基盤に係る国際標準化の取組に参画するとともに、基本的な機能は、既存のソフトウェアを活用して効率的に整備。
 - ①の基本機能は「Open Science Framework」（米国のデータ基盤）を構築したオープンソースソフトウェアを活用。
 - ②、③の基本機能は「JAIRO Cloud」や「CiNii」を構築したオープンソースソフトウェアを活用。

【研究データの保管・公開・検索の流れ】



※1 メタデータとは、データに関する基本情報（作成年月日、実験内容など）。これにより、データの利活用者にその内容を説明するとともに、検索による発見可能性を高めるために重要な情報となる。

※2 DOIとは、データに付与される国際的な識別子。公開されたデータへの永続的なアクセスを保障するために必要となる。