



オープンサイエンスの潮流を踏まえた 調査研究の開発

「科学技術・イノベーション政策の検討に資するエビデンス構築に向けて
—社会の大転換期に政策研究は何かできるか—」

2023年2月22日(水)

第15回政策研究レビューセミナー

文部科学省科学技術・学術政策研究所 (NISTEP)

データ解析政策研究室

室長 林 和弘

ICTの活用により知識をオープンにし，研究の加速化や新たな知識の創造などを促すオープンサイエンスの動きが活発化

- EU:**
- ヨーロピアン・オープン・サイエンス・クラウド (ESOC) が2018年に設立。
 - Horizon Europeの枠組みでは研究データはFAIR原則に基づき、「できるだけオープンに、必要に応じてクローズド」にし、また研究データマネージメントを必須とする。

- 豪州:**
- オーストラリア・リサーチ・データ・コモンズ (ARDC) を立ち上げ、研究データを中心とした情報基盤を構築中。

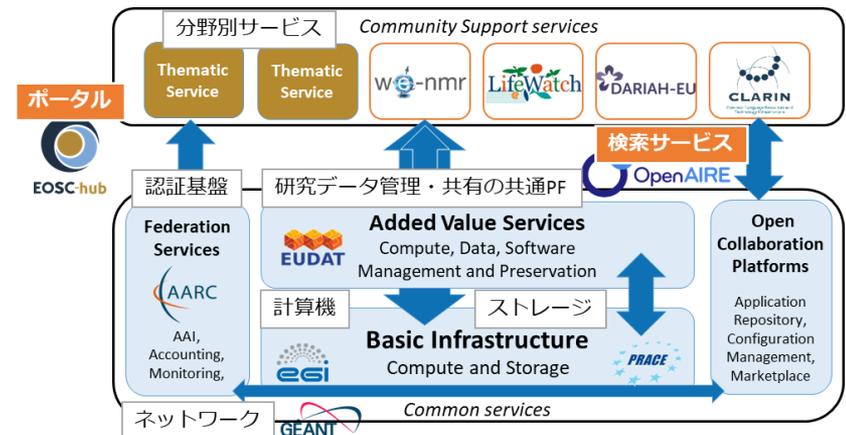
- 米国:**
- 米国国立標準技術研究所(NIST)は研究データフレームワークの設立を検討中。
 - OSTPはデータのアクセスや共有の在り方についてパブコメを実施。NIHがデータマネージメント&シェアリングポリシーを策定。

- G7:**
- オープンサイエンスWGを設置 (EUと日本が共同議長)。2016年より毎年WS等を開催。

- OECD:**
- 公的資金による研究データアクセスに関する理事会勧告 (2021年1月)

- UNESCO:**
- オープンサイエンス勧告 (2021年12月)

ヨーロッパ・オープン・サイエンス・クラウド



出典：国立情報学研究所 山地一禎教授 提供資料

林が
諮問委員等
専門家として貢献



G7 Research Compact

Open Science WG

As Open Societies with democratic values we believe in academic freedom. The freedom to pursue intellectual enquiry and to innovate allows us to make progress on shared issues and drive forward the frontiers of knowledge and discovery for the benefit of the entire world. We recognise that research and innovation are fundamentally global endeavours. Nations, citizens, institutions, and businesses have made huge strides forward, not otherwise possible, through open research collaboration across borders. Working together we will use our position as leading science nations to collaborate on global challenges, increase the transparency and integrity of research, and facilitate data free flow with trust to drive innovation and advance knowledge.

Shared Values

The global response to COVID-19 has demonstrated the progress that arises from long-term collaboration which puts science at the heart of prevention, preparedness, response, recovery and resilience. This progress requires sustained investment in research and supporting infrastructure, including in basic research and high-risk, high-reward undertakings. As our nations and communities start to recover from the pandemic and build resilience for future shocks, we will continue to work with our research and business communities to remove barriers to the open and rapid sharing of knowledge, data and tools, to the greatest extent possible, recognising the importance of research security in particular in cutting-edge fields, and to promote open science and increase open, safe and transparent dissemination of science to citizens, and to strive to minimise technology-related risk.

We can only tackle the greatest challenges that we face and will face over coming decades – such as climate change, pandemics and biodiversity loss – through transparent, open and agile research collaboration. We must bring the widest possible range of resources, expertise and perspectives to bear on solutions which will benefit people across the globe.

We commit to promoting international research cooperation and the conditions of freedom, independence, openness, reciprocity and transparency under which it flourishes. Our governments have the right and responsibility to effectively ensure the security and integrity of the research ecosystem, in partnership with the research community, preventing the theft, misuse and inappropriate exploitation of our intellectual property and personal data, and other forms of misconduct.

We are committed to developing a strong, diverse and resilient science and research community which is inclusive of all groups, as recognised by the Working Group on Financing Science for Inclusive Growth. It is important to deepen participation of underrepresented and marginalised communities and expand their participation in the research and innovation ecosystem. Inclusion will enhance the strength of our research base and increase momentum on dismantling the social, legal, and regulatory barriers limiting participation, and complementing our G7 gender equality goals by tackling gender gaps. Principles and practices of inclusive growth distribute the benefits of science among diverse communities and regions across the G7 and beyond.

UNESCO

193 Countries

UNESCO Recommendation on Open Science

At the 40th session of UNESCO's General Conference, 193 Members States tasked the Organization with the development of an international standard-setting instrument on Open Science in the form of a UNESCO Recommendation on Open Science to be adopted by Member States in 2021.

ISC (International Science Council)

Current

Open Science and the UNESCO initiative – opportunity to republish ISC statement

In this statement made by the ISC delegation to the UNESCO Special Committee meeting on Open Science, 6-12 May 2021, the delegation explores how the recommendation and potential cascading interventions by Member States could develop along two divergent pathways. ISC Members are invited to download the statement and republish on their websites and in their scientific journals.

support

UN

2nd Open Science Conference

FROM TACKLING THE PANDEMIC TO ADDRESSING CLIMATE CHANGE

Organized by the United Nations Dag Hammarskjöld Library and the United Nations Department of Economic and Social Affairs

Open Science Conference 2021

support

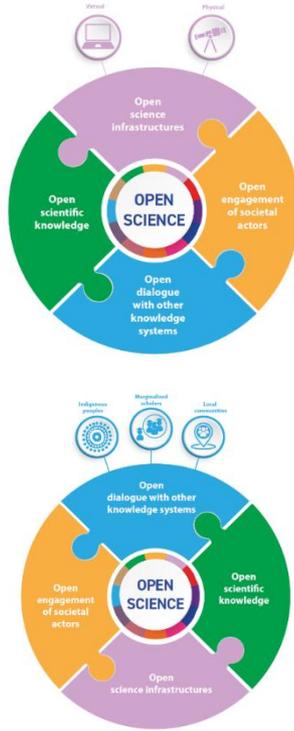
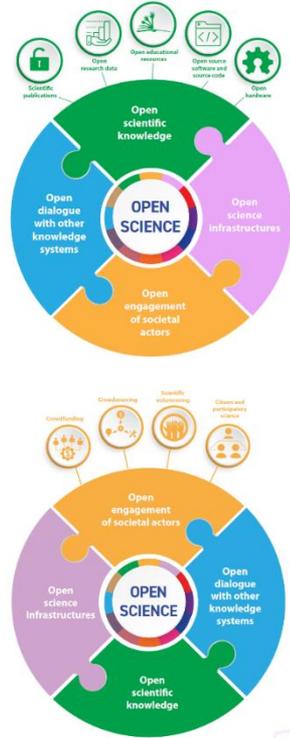
<https://www.mofa.go.jp/mofaj/files/100200013.pdf>
<https://en.unesco.org/science-sustainable-future/open-science/recommendation>
<https://council.science/current/news/open-science-and-the-unesco-initiative/>
<https://www.un.org/en/library/OS21>

新しいイノベーション基盤に

先進国としてのリーダーシップ
(研究力、産業)

格差社会の解消、新しい知の営み

国際社会のリーダーシップ
(外交、教育)



UNESCO勧告より

- より開かれた科学知識
 - ◆ 論文以外の様々な成果もオープンに
- オープンサイエンス基盤（インフラ）
 - ◆ 人と機械が読めるインフラ整備
- 社会的アクターのオープンな関与
 - ◆ 市民の参画による新しい研究スタイル
- 他の知識システムとの開かれた対話
 - ◆ 先住民や地域が持つ伝統的な知識の導入と活用



背景

- オープンサイエンスは社会変動の駆動力の一つ
 - ◆ UNESCO,G7等の主要アジェンダ：情報格差解消やイノベーションを加速させるものとして注目
- オープンサイエンス政策は包括的な科学，社会，科学と社会の変容を見越し，研究成果（知識）の幅広い波及に着目
 - ◆ 研究データの共有に重点がおかれつつも、2022年度は論文のオープンアクセス（OA）も再注目を浴びた（米国OSTP指令等）
- 実際の研究のあり方，研究成果自体も変えつつある状況をいかに定量的に把握するか
 - ◆ 変容のシグナルをつかむための定期的なモニタリングの重要性
 - ◆ これからの変容に関しては指標化を目指す前段階のモデル化が求められる段階

今回の発表概要

- A) オープンアクセス（OA）とプレプリントの利活用と認識に関する調査2022
 - ◆ 分野別を含め、OAとプレプリントの受容と利活用の実態を調査
- B) 研究活動におけるオープンソース・データの利用に関する簡易調査
 - ◆ プレプリントに言及されているオープンソース・データの言及からその利活用状況を把握



A OAとプレプリントの利活用と認識に関する調査2022

<https://doi.org/10.15108/rmxxx>

1. 調査方法と対象

■ 調査方法 ``オンライン質問紙調査 (Cuenote)

■ 調査対象 ``NISTEP科学技術専門科ネットワーク

- ✓ 紹介によって選出される産学官の研究者、技術者、マネージャーなど 約2,000名の専門家集団
- ✓ 多分野かつ幅広い年齢層の回答者が含まれる

<https://www.nistep.go.jp/activities/st-experts-network>

■ 調査期間 ``2022年7月19日~8月22日

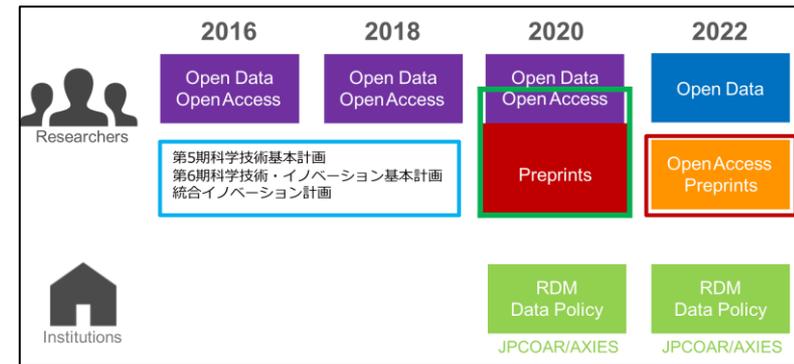
- ✓ 予備調査 ``2022年6月7日~6月20日
- ✓ リマインダ ``未回答者を対象として7月27日、8月1日に送信

■ 依頼方法 ``E-mail

- ✓ 多重回答を防ぐために個別URLを設定
- ✓ 回答は1回のみ可能

■ 回答率 ``70.2%

- ✓ 依頼 ``1,671名
- ✓ 有効回答 ``1,173名



2年おきに、調査、内容や対象を拡大中

回答者の概要 (n=1,173)

研究分野



所属

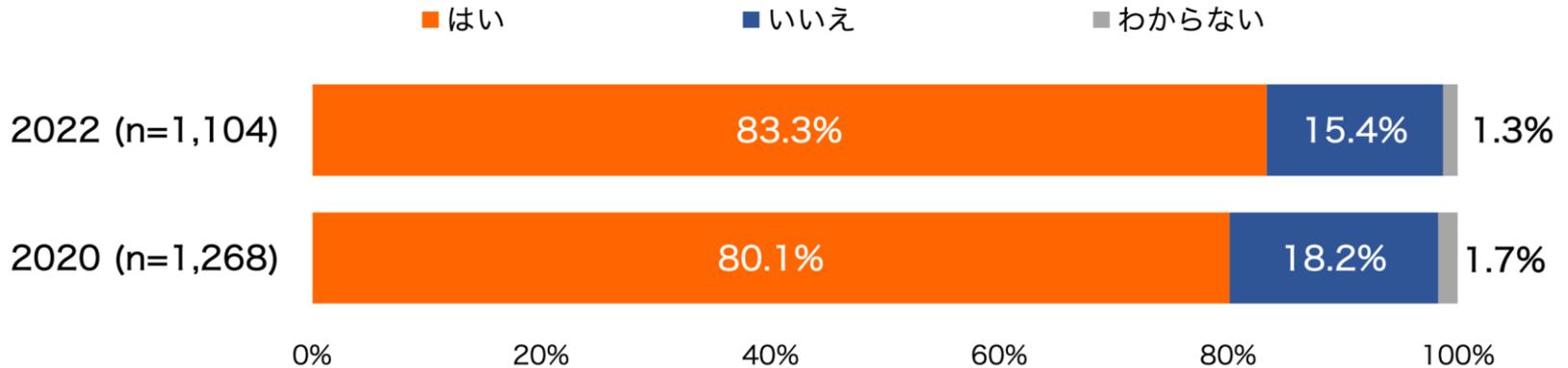


その他 0.1%

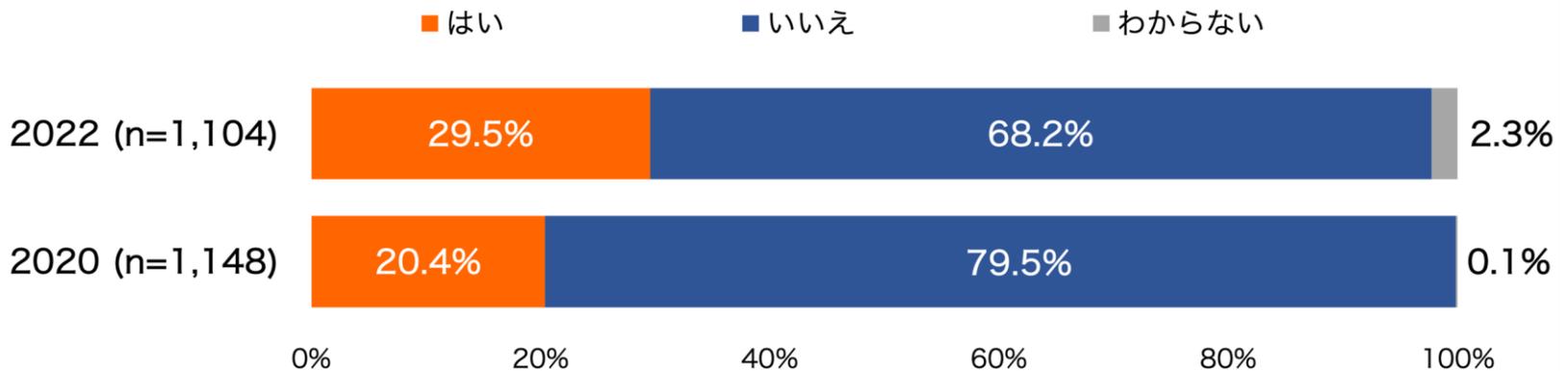
年齢層

年代	回答者数	比率
30代以下	256	21.8%
40代	600	51.2%
50代	241	20.5%
60代以上	76	6.5%
合計	1,173	100.0%

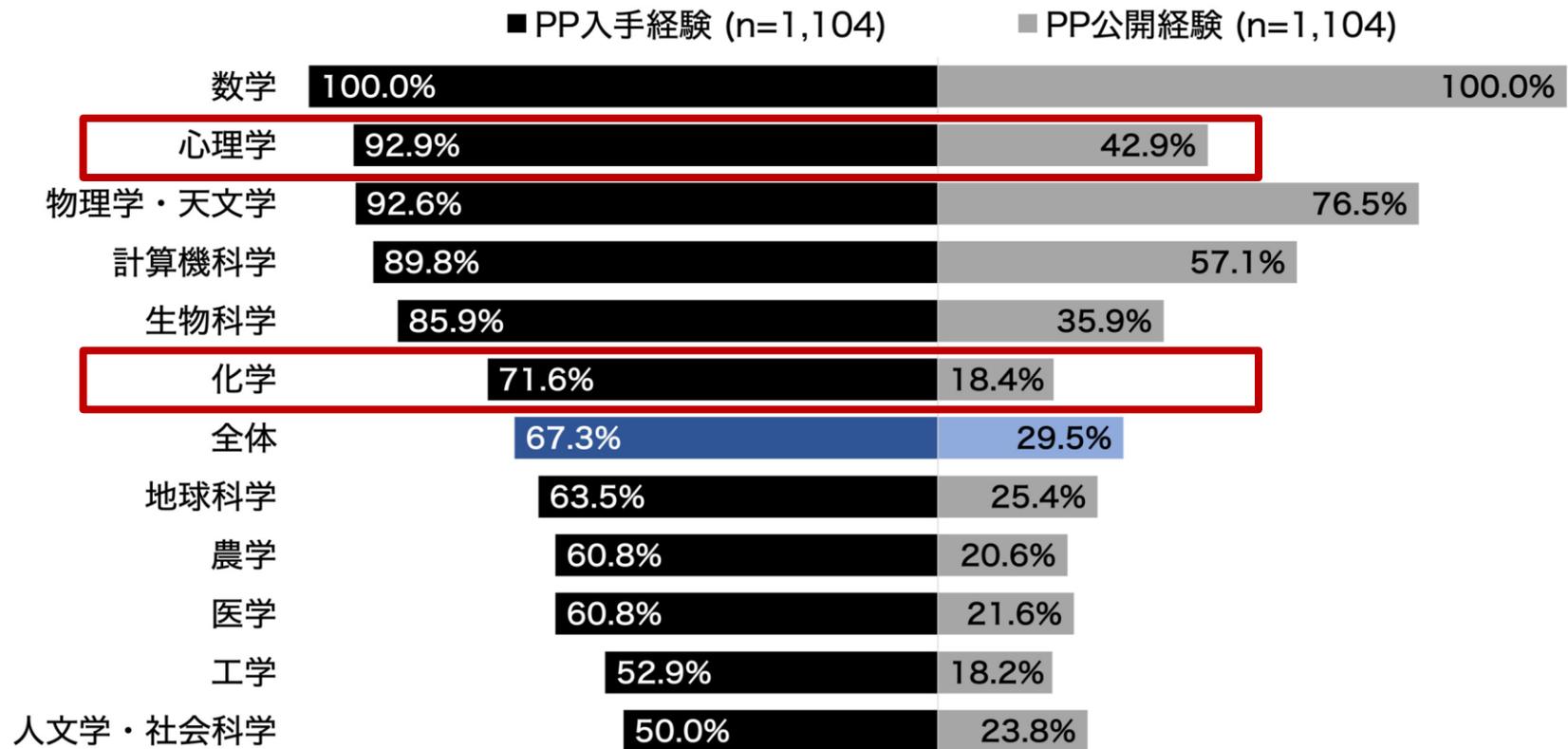
論文



PP



分野別：PPの入手経験と公開経験



Pearsonの相関係数 0.83, $p < 0.01$

	論文	プレプリント
2022 /20	<ul style="list-style-type: none"> 1 資金がないから 2 投稿したい雑誌がOAではないから 	<ul style="list-style-type: none"> 1 PPを公開する必要性を感じないから 2 最初に査読誌に投稿したいから
2020 から 増加	<ul style="list-style-type: none"> OAではない発表方式の方が望ましいと思うから 時間がないから 特に理由はない リポジトリなどの公開手段がないから 	<ul style="list-style-type: none"> 業績にならないから

黒字＝内的要因、赤字＝外的要因

出版形態	選択率が高い分野	選択率が低い分野
プレプリント**	生物科学、数学、物理学・天文学	工学
書籍***	HSS(人文学・社会科学)	医学、生物科学
査読付き論文／レビュー論文*		工学、HSS
査読付き会議録***	計算機科学、工学、物理学・天文学	生物科学、地球科学、HSS、医学、化学
テクニカルレポート	農学、工学	生物科学、化学
研究データ***	地球科学	
プログラムコード、ソフトウェア***	計算機科学、地球科学	
特許***	工学、農学	HSS、物理学・天文学、生物科学、数学、地球科学

$p < .05$ *, $p < .01$ **, $p < .001$ ***

2023年3月に調査資料として公開予定



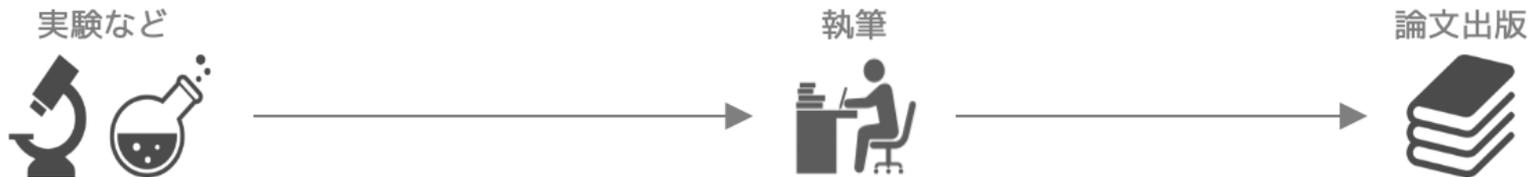
B 研究活動におけるオープンソース・データの利用に関する簡易調査 [調査資料 No.324]

<https://doi.org/10.15108/rm324>

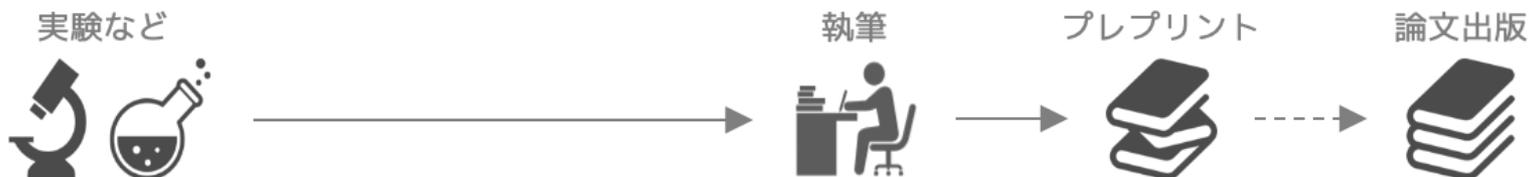
1. オープンサイエンス・DXによる研究プロセスと成果公開の変化

一般的な自然科学系の研究成果公開プロセス

【但し、どのプロセスも執筆を経て実験を見直すなど、実際には決してリニアではない】

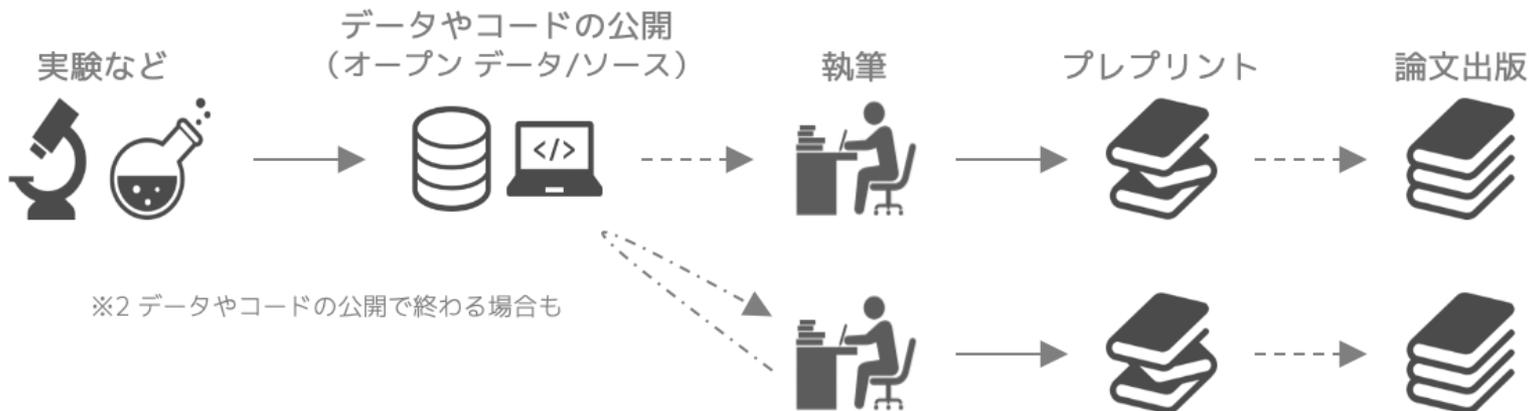


プレプリントを活用した研究成果公開プロセス



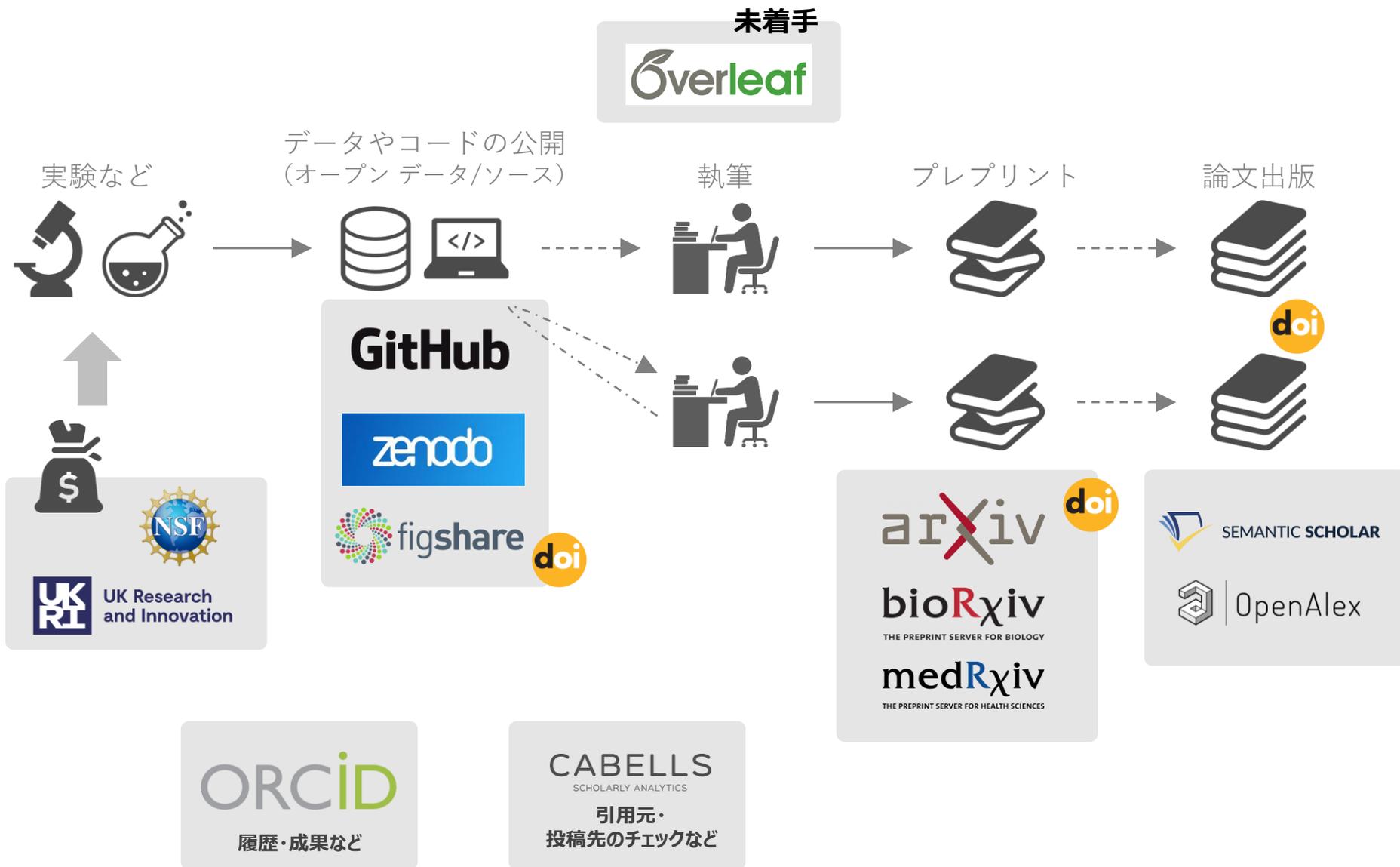
※1 プレプリント公開で終わる場合 / プレプリントを経ず出版の場合も

DXにより予見される自然科学研究のプロセス



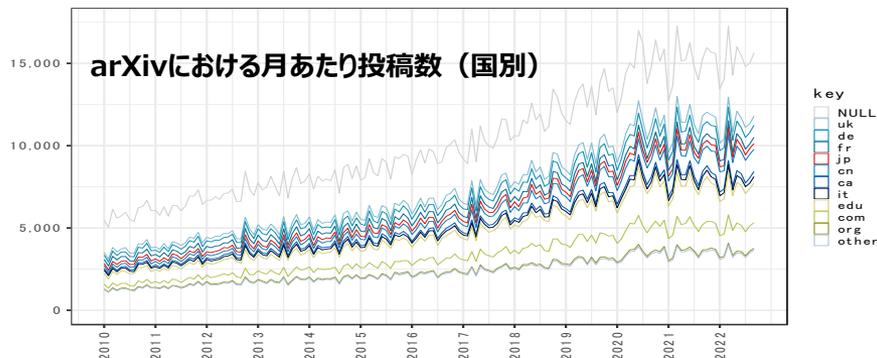
※2 データやコードの公開で終わる場合も

※3 オープンなデータ等を使うことで、実験を行わない場合も

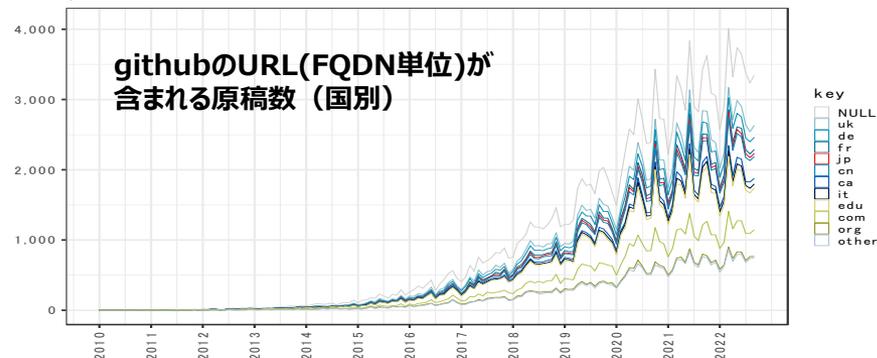


- 研究成果における OSS※を含む、オープンデータ利用度合いの定量把握を目的とした分析
- 利用度合いに対する期待およびデータ取得の容易性から、物理・情報系のプレプリントサーバーarXivのデータを対象
 - ・ オープンデータ以前に、論文中に何らかのDOIを記載するものの割合は徐々に伸びている
 - ・ 分野特性とあいまって OSS(github) に関する言及は全体の2割を超えるなど、既存研究の分析とも一致する傾向 (国別は本報の差分)
 - ・ オープンデータ (Zenodo, Figshare) の利用については、未だ黎明期にあり、引き続きの動向観察が期待される

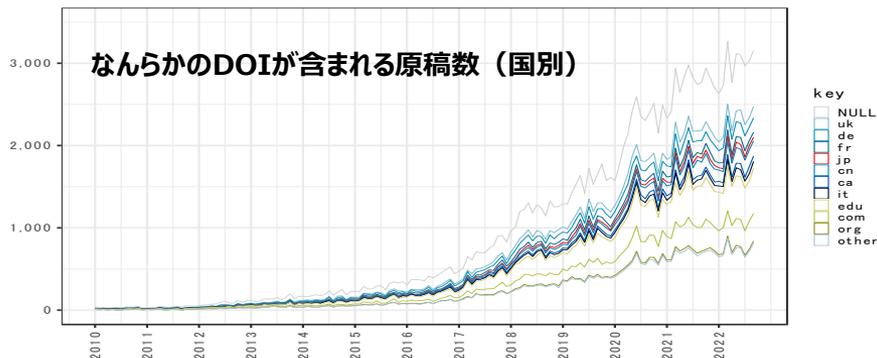
General Counts



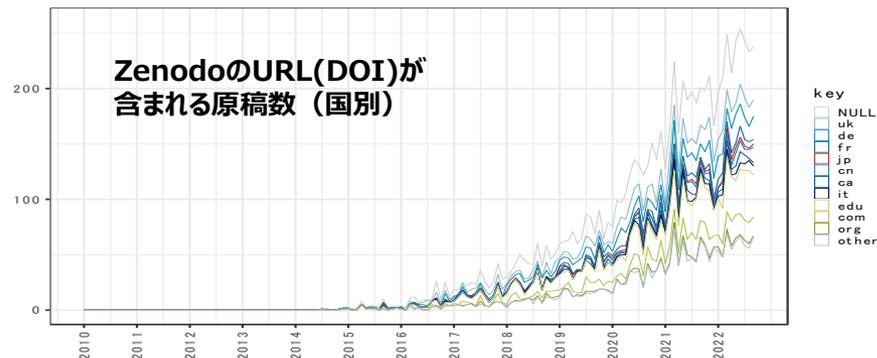
github Counts



DOI Counts



zenodo Counts



【政策等への示唆】

- ・ 第6期STI基本計画に記載のある、DXによる研究活動の変化に関する分析手法・指標になる可能性
- ・ 日本に着目した場合、DOIの普及割合に対して OSSの利用度合いが低く、他国とは異なる傾向を有する
- ・ オープンデータ利用が進むことは論を待たないが、プラットフォームはまだ発展途上にあり、今度の動向には要注目

NISTEP実態調査2016, 2018, 2020

■ NISTEP科学技術専門家ネットワークを対象としたデータ公開を中心としたオープンサイエンスの実態や課題を把握するための調査

- ◆ データ公開と（比較として）OA*1論文の現状
- ◆ データ公開の障壁
- ◆ 公開データの利用状況
- ◆ DMP*2の作成状況（2018から）

*1 OA: Open Access
*2 DMP: Data Management Plan

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。文部科学省科学技術・学術政策研究所, 2017, NISTEP RESEARCH MATERIAL No.268, <https://doi.org/10.15108/rm268> (※ 0巻以前の成果)

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。文部科学省科学技術・学術政策研究所, 2020, NISTEP RESEARCH MATERIAL No.289, <https://doi.org/10.15108/rm289> (※ 0巻以前の成果)

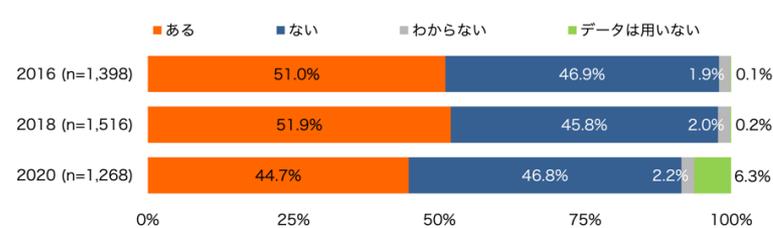
研究データ公開と論文のオープンアクセスに関する実態調査2020, NISTEP RESEARCH MATERIAL, 2021, No.316, <https://doi.org/10.15108/rm316>

2017年12月

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

文部科学省 科学技術・学術政策研究所
科学技術専門センター
池内 智哉、林 和弘、赤澤 伸一

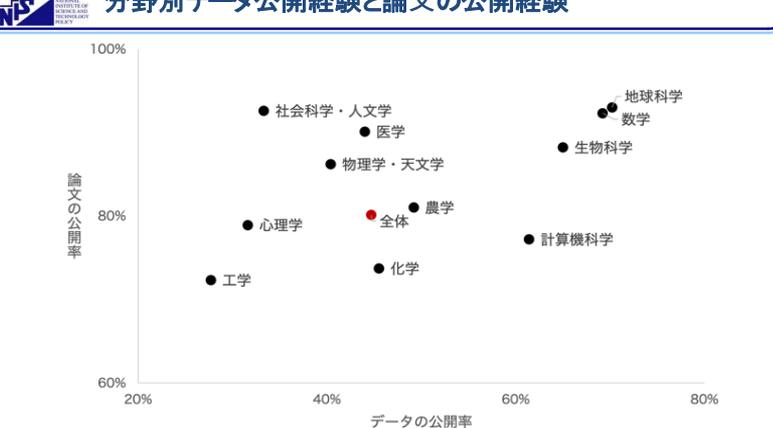
データの公開経験



年	あり	なし	わからない	データは用いない
2016 (n=1,398)	51.0%	46.9%	1.9%	0.1%
2018 (n=1,516)	51.9%	45.8%	2.0%	0.2%
2020 (n=1,268)	44.7%	46.8%	2.2%	6.3%

- 2016/2018年調査と比較して、公開率がやや低下
- 低下の要因
 - ◆ アンケートシステムの変更に伴う質問方法が変更されたことによるもの
 - ◆ 回答者の揺れ(同一回答者の、同一システム・同一質問の2016/2018年調査でも矛盾あり、データ公開以外の項目も矛盾あり)

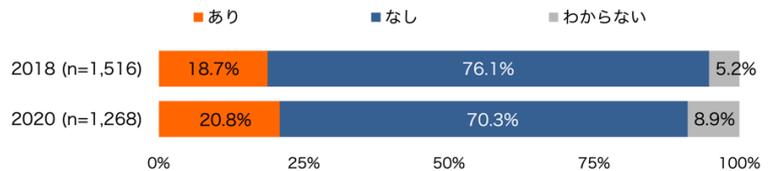
分野別データ公開経験と論文の公開経験



分野	データの公開率 (x)	論文の公開率 (y)
地球科学	~75%	~95%
数学	~70%	~90%
生物学	~65%	~85%
社会科学・人文学	~35%	~90%
医学	~45%	~85%
物理学・天文学	~40%	~80%
農学	~50%	~80%
心理学	~30%	~75%
全体	~45%	~75%
工学	~25%	~65%
化学	~45%	~65%
計算機科学	~60%	~75%

- データ公開経験とOA経験に有意な相関はなし
- いずれも高いのは地球科学・数学、いずれも低いのは工学
- OA率が高いのは社会科学・人文学、データ公開率が高いのはCS

データマネジメントプラン(DMP)の作成経験

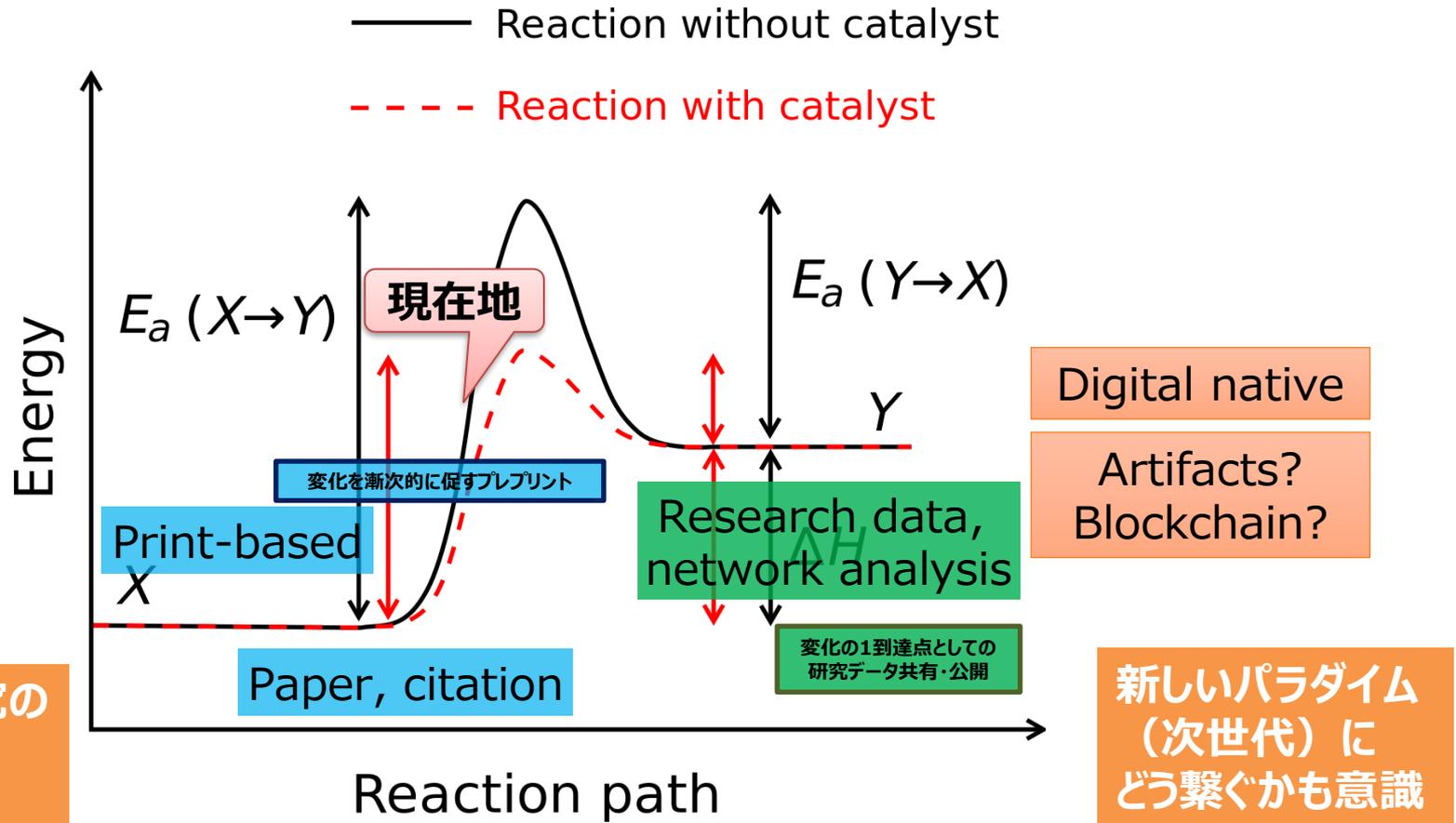


年	あり	なし	わからない
2018 (n=1,516)	18.7%	76.1%	5.2%
2020 (n=1,268)	20.8%	70.3%	8.9%

- 作成経験をもつ研究者の比率がわずかに増加
- 「わからない」の選択率が増加

- オープンサイエンス政策は包括的な科学と社会の変容を見越し、研究成果（知識）の幅広い波及に着目
 - ◆ 具体の政策としては、論文や研究データの利活用促進による科学技術・学術の進展とイノベーションの加速に注目

- この潮流を踏まえ、NISTEPでは、研究論文（OA）、プレプリント、研究データの利活用に関する調査を実施し、変容のモデル化や指標化に取り組んでいる
 - A) オープンアクセス（OA）とプレプリントの利活用と認識に関する調査2022
 - ✓ OAとプレプリントの受容と利活用は進展
 - ✓ 分野別に比較的大きな差が存在しており、プレプリントを業績として認める動きのある分野も
 - ✓ OAとプレプリントそれぞれの公開阻害要因なども明らかに
 - B) 研究活動におけるオープンソース・データの利用に関する簡易調査
 - ✓ オープンデータ以前に、論文中に何らかのDOIを記載するものの割合は徐々に伸びている
 - ✓ 分野特性とあいまって OSS(github) に関する言及は全体の2割を超えるなど、既存研究の分析とも一致する傾向
 - ✓ オープンデータ (Zenodo, Figshare) の利用については、増加中ではあるが未だ黎明期にあり、引き続きの動向観察が期待される



これまでの研究の
エコシステムは
依然重要

(5-10年の計)

- 論文やプレプリントを調査することで漸次的な変化を捉え
- 研究データ共有・公開を調査することで、非連続な変化を捉える

(100年の計)

(参考) オープンサイエンス政策への貢献と実践

オープンサイエンスがもたらす科学と社会の変容を政策と実践の両面から促し、モニタリングする調査研究



国際的な活動への貢献
調査、ガイドライン作りへの協力

専門家派遣



G7 オープンサイエンスWGワークショップ
(2019.6)

共同議長国の専門家としてとして
企画運営をサポート



多様な分野やセク
ターとの対話



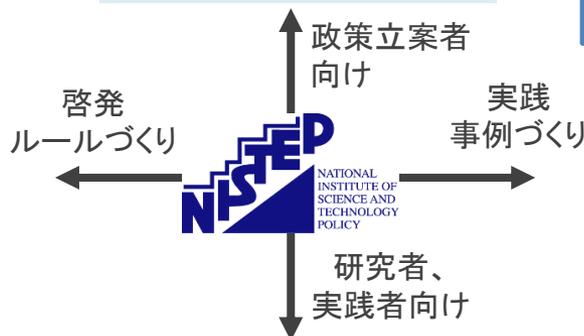
J P C O A R
オープンアクセスリポジトリ推進協会

学術関連コミュニティとの対話
政策への橋渡し



国内政策作りのサポート
エビデンス提供

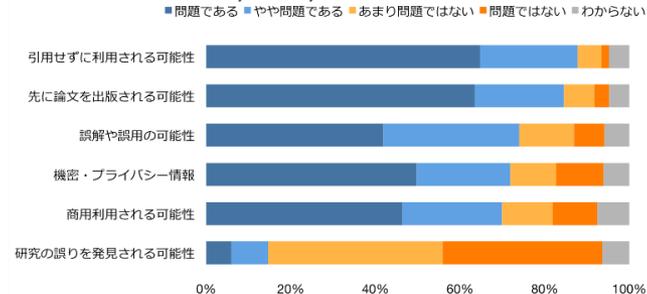
科学技術基本計画および統合イ
ノベーション戦略策定への協力



NISTEP共催イベントの開催等

国内啓発活動
シチズンサイエンス

データを公開する場合の懸念の強さ
(n=1,396, 2016調査)



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
<https://doi.org/10.15108/rm268>



RDA (研究データ
連盟) の当該IGの
共同議長

国際的なデータ共有の
共通質問紙の検討

研究データ活用協議会



Research Data Utilization Forum

研究データリポジトリ
に関する小委員会ガ
イドラインを内閣府
版の草案に

草の根活動への参画と
政策への橋渡し

I 勧告の目的

オープンサイエンスにおける学術分野や地域の違いや、科学者やその他関係者における学問の自由及びジェンダー変革的なアプローチ、その他特別な課題を考慮した、国内外のデジタル、技術及び知識の格差縮小に貢献するオープンサイエンスの政策や実践のための国際的な枠組みの提供。

II オープンサイエンスの定義

多言語の科学的知識を誰もがオープンに入手しアクセスし再利用することを可能にし、科学と社会の利益のために科学的協力や情報共有を増大させ、従来の科学コミュニティを超えて、科学的知識の創造、評価、伝達のプロセスを関係者に開放することを目的とした、様々な行動や実践を組み合わせた包摂的な概念。

III オープンサイエンスの中核的な価値

- ① 質及び誠実性、② 集団的利益、③ 公平と公正、④ 多様性と包摂性

IV 行動の分野

- ① 利点と課題を伴うオープンサイエンス及びオープンサイエンスへの多様な道筋についての共通理解の促進
- ② オープンサイエンスを可能にする政策環境の整備
- ③ オープンサイエンスのインフラとサービスへの投資
- ④ オープンサイエンスのための人的資源、トレーニング、教育、デジタルリテラシー、能力開発への投資
- ⑤ オープンサイエンスの文化の醸成、オープンサイエンスへのインセンティブの調整
- ⑥ 科学的プロセスの各段階におけるオープンサイエンスのための革新的なアプローチの促進
- ⑦ オープンサイエンスの文脈におけるデジタル、技術、知識の格差縮小のための、国際的かつ多面的な利害関係者との協力の促進

V モニタリング

加盟国は、その個別の事情やガバナンス構造及び憲法上の規定に従い、定量的および定性的なアプローチを適宜組み合わせ、オープンサイエンスに関連する政策及び仕組みをモニタリングすべきである。

(上記勧告に法的拘束力はなし)



- 我が国では、第5期科学技術基本計画より「オープンサイエンスの推進」を謳っており、第6期計画においてもデータポリシーの策定状況等具体的な目標設定を掲げ、イノベーション基盤整備の重要な項目としてオープンサイエンスの積極的な推進を取り入れている。
- また、公的資金による研究データの管理・利活用に関する基本的な考え方(国のデータポリシー)において、メタデータ共通項目の設定や研究データ基盤システム(NII RDC)を研究データ管理・利活用の中核的なプラットフォームとして定めている。