

SPARC Japan ニュースレターでは、各回セミナーの報告に講演やパネルディスカッションを書き起こしたドキュメントを加え、さらにそのほかの SPARC Japan の活動をご紹介します。

※所属、肩書はすべて開催当時のものです。

CONTENTS

■ SPARC Japan 活動報告

学術情報流通推進委員会
学術情報流通推進委員会の終了
「SPARC Japan 活動の振り返りと今後の方向性」の公開
arXiv.org 活動支援について
CLOCKSS 活動支援について
SCOAP³ 活動支援について

■ SPARC Japan セミナー報告

企画概要
参加者から
企画後記
ドキュメント
(講演・総合討論)

■ SPARC Japan 活動報告

学術情報流通推進委員会

学術情報流通推進委員会の会議資料をウェブサイトで公開しています。

<https://www.nii.ac.jp/sparc/about/committee/>



学術情報流通推進委員会の終了

SPARC Japan は、2003～2018 年度は「国際学術情報流通基盤整備事業」として、2019～2021 年度は「学術情報流通推進委員会」として活動し、国内学会誌の電子ジャーナル化、オープンアクセス・オープンサイエンスのアドボカシー活動、国際イニシアティブ（arXiv.org、CLOCKSS、SCOAP³）との連携等に一定の役割を果たしてきました。

その結果、大学図書館界をはじめとしたステークホルダーにおけるオープンアクセス・オープンサイエンスに係る、各種会議体やアドボカシー活動を行う団体の活動が拡充しており、検討の結果、学術情報流通推進委員会は、第 1 期（2019～2021 年度）をもって活動を終了することにしました。

なお、以下は継続して活動します。

- ・オープンアクセス・オープンサイエンス等のアドボカシー活動（SPARC Japan セミナー）
- ・国際イニシアティブ（arXiv.org、CLOCKSS、SCOAP³）との連携

「SPARC Japan 活動の振り返りと今後の方向性」の公開

学術情報流通推進委員会の終了に伴い、これまでの SPARC Japan の活動を振り返り、関係するステークホルダーに要請される今後の活動の方向性を提言した報告書「SPARC Japan 活動の振り返りと今後の方向性」を作成しました。

報告書は以下ウェブサイトで開催しています。

<https://doi.org/10.20736/0002000313>

arXiv.org 活動支援について

arXiv.org < <https://arxiv.org/> > は、物理学、数学、コンピュータサイエンス等のプレプリントサーバとして、同分野を中心とした研究成果のオープンアクセスとアーカイブを進めています。2022 年 1 月に収録論文数が 200 万件を突破し、年間ダウンロード数は約 3.7 億件以上で、総ダウンロード数は 2022 年 4 月時点で約 23 億件です。

利用件数上位の機関による財政支援があり、2013 年に開始した「arXiv.org 会員制プログラム」により 2019 年末で 30 カ国 243 機関が参加しています。

日本においては、arXiv.org の利用実績が多い研究機関がコンソーシアムを形成してこの取組みを支援しており、2009 年から国立情報学研究所にて、参加意向確認および年会費の取りまとめを行っています。

arXiv.org の利用頻度が高い国内機関に対して 2021 年の参加意向調査を行った結果、2022 年 3 月末現在の参加機関は 16 機関です。

2022 年 1 月より新規記事に対して DOI が自動的に付与されることとなりました。

arXiv.org に関しては以下ウェブサイトを公開しています。

<https://www.nii.ac.jp/sparc/about/international/arXiv/>

関連会議は、以下のとおり開催され、日本の参加機関を代表して国立情報学研究所武田英明教授が参加しました。

- ・ Member Advisory Board Annual Meeting 2021 年 10 月 4 日～5 日（オンライン）

CLOCKSS 活動支援について

CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) < <https://clockss.org/> > は、全世界の研究者のために電子的な学術コンテンツの長期保存を実現することを目的として活動しています。具体的にはアーカイブとそれを運営するコミュニティを構築して、コンテンツが出版社から提供されなくなった場合にアーカイブ上のコンテンツを広く利用できるようにするなどの取り組みを行っています。また、コンテンツを保持する世界 12 のノード機関の一つとして、国立情報学研究所が参画しています。

日本においては、2013 年から国立情報学研究所にて、参加意向確認および年会費の取りまとめを行っています。

大学図書館コンソーシアム連合 (JUSTICE) 会員館を中心に 2021 年の参加意向調査を行った結果、2022 年 3 月末現在の参加機関は 101 機関です。

関連会議は、以下のとおり開催され、日本の参加機関を代表して国立情報学研究所武田英明教授が参加しました。

・CLOCKSS Board of Directors 会議 2021年6月23日、9月28日、2022年2月22日（オンライン）

SCOAP³ 活動支援について

SCOAP³<<https://scoap3.org/>>は、欧州原子核研究機構（CERN）が主導する、高エネルギー物理学（HEP）分野の査読付き学術雑誌論文のオープンアクセスを実現する国際連携プロジェクトです。日本においては、SCOAP³対象誌を購読している研究機関を中心に、従来図書館が出版社に支払ってきた購読料を、論文出版加工料（APC）に振り替えることで、HEP分野の主要な学術雑誌のオープンアクセスを支援しています。

日本においては、2014年から国立情報学研究所にて、参加意向確認および年会費の取りまとめを行っています。

フェーズ3（2020-2022年）を支援する国内機関を中心に2021年の参加意向調査を行った結果、2022年3月末現在の参加機関は82機関です。

また、2019年のGOVERNING COUNCIL会議において1年限りのプロジェクトとして承認された書籍を対象とするSCOAP³ for Booksパイロットプロジェクト<<https://scoap3.org/scoap3-books/>>について、2021年3月にSCOAP³参加機関宛に関心調査を行った結果、18機関に賛同をいただきました。

SCOAP³に関しては以下ウェブサイトを公開しています。

<https://www.nii.ac.jp/sparc/about/international/scoap3/>

関連会議は、以下のとおり開催され、日本の参加機関を代表して高エネルギー加速器研究機構（KEK）池松克昌特任専任 URA および国立情報学研究所山地一禎教授が参加しました。

・GOVERNING COUNCIL 会議 2021年11月3日～4日（オンライン）

■ SPARC Japan セミナー報告

SPARC Japan セミナー2021



「研究データポリシーが目指すものとは」

2022年2月22日（火） オンライン開催 参加者（Webex Events）：297名

次ページ以降に、当日参加者のコメント（抜粋）、企画後記およびドキュメント全文（再掲）を掲載しています。その他の情報は SPARC Japan の Web サイトをご覧ください。

<https://www.nii.ac.jp/sparc/event/2021/20220222.html>

企画概要



2021年にユネスコよりオープンサイエンスの勧告が出されて、オープンサイエンスと研究成果の共有・公開が国際的に最重要視される中、日本ではその具体化を進め、統合イノベーション戦略2021により、大学・大学共同利用機関法人・国立研究開発法人等の機関には、データポリシーの策定が年限付きで求められています。

このため各機関では、データポリシーの検討が進んでいますが、その過程において、機関の置かれた環境に応じて様々な課題が見え始めています。

その一方、データポリシーの策定はオープンサイエンス実現のための手段であり、それ自体が目標ではありません。

本セミナーでは、政策関係者・大学関係者・研究データ流通関係者など、様々なステークホルダーが一堂に集い、データポリシーの意義と、その先のビジョンの一つである FAIR 原則に基づくデータの利活用推進のために必要な考えやアクションは何かについて討論を交わし、地に足のついた研究データポリシーの策定に役立ちます。



参加者から

(大学/図書館関係)

・研究データポリシーの策定そのものではなく、それを含んだ研究データ活用全体の在り方について、様々な機関や分野の関係者から意見を聴くことができ、自機関での研究データ管理を考える参考になった。

・研究データ管理には管理の問題と利活用の問題があり、更に前者には研究者の立場から評価問題や、後者には政策的な「国益」の思惑と研究機関としてよりよい研究をしたいという展望とがあることが実感としてわかった。

・研究データポリシー策定の期限が迫る中で、どのような点を考えて策定していくべきか、など大変参考になりました。また、現在策定中やこれから策定に向けて動く大学、機関様も多いと思われ、機関の中で他部署との連携の必要性を再確認できました。

・大変参考になりました。ポリシーを策定する前に自分の大学の研究の特色、ポリシーは何のために作るのか、研究活動の方向性を明確にするという話が非常に参考になりました。

(大学/研究者)

・データポリシーの策定は「機関のミッションに立ち戻る機会です」という発言がよかった！

・2時間議論の予定で驚きましたが、とてもすばらしく、あっという間でした。ありがとうございました。

・今後のデータポリシーをどうすべきなのか、大変参考になりました。一定期間後の開催を希望します。

(大学/その他)

・色々なステークホルダーの具体的な問題点が把握できてよかったです。ありがとうございました。

(その他/その他)

・国の状況と大学等実際に研究データの利活用を担う方の意見を聞いたのが良かった。

・文科省担当官による政策紹介と展望があったので、制度整備のスケジュールを含む全容の俯瞰ができてよかった。

企画後記



😊 これから策定されるデータポリシーが、研究者に負担を強いるものではなく、データの利活用を推進し、科学を進展させるものにするためにはどうすればよいのか、何が課題なのか、たくさんのヒントを得られました。引き続き、国内外の動向をふまえた議論を定期的に行えればと思います。

池内 有為
(文教大学)

😊 セミナー企画WG委員として、今回始めて関わらせていただいたテーマが「研究データポリシーが目指すものとは」でした。一大学図書館員として所属機関で関わるテーマでもあり、田野学長のお考えを踏まえ、方向性や課題を確認する機会にもなりました。多彩な登壇者からの多角的な意見交換もあり、

同じく研究データに関わる方々に学びになるセミナーとなったのではないかと思います。

上野 友稔
(電気通信大学)

😊 今年度のSPARC Japanセミナーは諸事情で年1回の開催となりました。そのため、最初は色々詰め込んだ総花的な企画となりましたが、研究データとポリシーに絞ることで時間も内容も(これでも)コンパクトになりました。多くの関心をいただけてホッとしております。

林 和弘
(科学技術・学術政策研究所)

😊 2 部にわたる総合討論で様々な立場の関係者の方々の見解をうかがうことができ、非常に充実したセミナーになったと思います。企画 WG メンバーとしての参加は今回が最後となりますが、企画から携わせていただくことで、非常に多くのことを学ばせていただきました。ありがとうございました。

安原 通代
(国立情報学研究所)

😊 本年度はセミナー回数が1回だったので、WG 全メンバーでセミナー内容を検討しました。これは初めての試みで、何度も会合を重ねるなどいつもの年より時間はかかりましたが、WG メンバー間の理解を深めることができました。セミナー当日に交わされた、深くかつ射程の広い議論は、こうした WG 内の議論があつてこそ可能になったと思います。

八塚 茂
(バイオサイエンスデータベースセンター)

😊 今回のセミナーでは、研究データポリシーの策定について様々な立場からの議論やコメントをいただきました。研究データポリシーは、よりよい研究活動が行われるために作られるものであって、そのためには単に政策上の要請があるからということにとどまることなく、丁寧な議論が積み重ねられ、前向きに取り組まれる必要があると思います。今回のセミナーもそういった議論の場の一つになったのではないかと思います。

矢吹 命大
(横浜国立大学)

😊 討論だけでも 2 時間に及ぶ長丁場！ しかもオンライン開催にも関わらず、終始、聴衆の皆様の高い関心と熱気が伝わってくるようでした。今回提示された多様なご意見・ご経験が、各機関での研究データポリシー策定の一助となれば幸いです。

山形 知実
(北海道大学)

本誌についてのお問い合わせ

国立情報学研究所 SPARC 担当

E-mail co_sparc_all@nii.ac.jp FAX 03-4212-2375

<https://www.nii.ac.jp/sparc/>

SPARC Japan セミナー2021

「研究データポリシーが目指すものとは」

開会挨拶 / 概要説明

池内 有為

(文教大学 文学部)

池内 有為



文教大学文学部専任講師。博士（図書館情報学）。慶應義塾大学法学部卒業（1995年）、同文学研究科図書館・情報学修士課程修了（1997年）の後、1997年から2005年までフェリス女学院大学附属図書館勤務。主婦、筑波大学大学院博士後期課程を経て、2019年より現職。専門領域は研究データ共有とオープンサイエンス。日本学術振興会人文・社会科学データインフラストラクチャー構築推進センター研究員、文部科学省科学技術・学術政策研究所客員研究官、2021年度SPARC Japanセミナー企画ワーキングメンバー。

本セミナーの概要

2021年にユネスコからオープンサイエンスの勧告が出され、オープンサイエンスおよび研究成果の共有・公開が国際的に進められてまいりました。日本においても統合イノベーション戦略 2021 によって、大学などの研究機関にはデータポリシーの策定が年限付きで求められています。このため、各機関ではデータポリシーの検討が進んでいることと思いますが、その過程において、機関の置かれた環境に応じてさまざまな課題が見え始めています。その一方で、データポリシーの策定はオープンサイエンス実現のための手段であって、それ自体が目標というわけではありません。

そこで本セミナーでは、政策関係者・大学関係者・研究データ流通関係者など、さまざまなステークホルダーが一堂に集い、データポリシーの意義と、その先のビジョンの一つである FAIR 原則に基づくデータの利活用推進のために必要な考え方や具体的なアクションは何かということについて、ざっくばらんに討論を交わし、地に足のついた研究データポリシーの策定に役立てていきたいと考えております。

本セミナーの構成

本セミナーでは、まず3名の講演者の方からご講演いただきます。初めに文部科学省研究振興局の三宅様から「研究 DX を巡る政策動向から見る研究データポリシーの役割」について、続いて、電気通信大学の田野学長から「大学が目指す共創進化スマート社会とその Science2.0 への展開」について、最後に NII の大波様から「学術情報インフラが実現する研究データの管理と循環」と題して NII 研究データ基盤（NII Research Data Cloud）についてお話しいたします。

その後、総合討論第1部「研究データポリシーをめぐる論点整理を中心に」では、データポリシーが目指す理念、科学研究をどのようにしていくのかということについてディスカッションします。総合討論第2部「研究データに関わる各ステークホルダーとの議論」では、ご講演者3名に加えて、既にデータポリシーを策定された研究機関の方、さらに SPARC Japan セミナーの企画ワーキンググループのメンバーも加わり、ポリシーの理念はよく分かるけれども、実際にやってみてどうなのか、現場として実務の上ではどんなギャッ

ブがあるのか、内部で運用した際にどのような大変なことがあるのか、そして、どんなところが肝となっているのかという論点について意見を交換できればと思っています。

SPARC Japan は、講演者の方だけではなく、会場の皆さまを加えてのディスカッションが肝だと思っています。積極的なご質問やご発言を期待しております。

SPARC Japan セミナー2021

「研究データポリシーが目指すものとは」

研究 DX を巡る政策動向から見る 研究データポリシーの役割

三宅 隆悟

(文部科学省 学術基盤整備室)

講演要旨



第6期科学技術・イノベーション基本計画では、オープンサイエンスとデータ駆動型研究の推進が柱の一つとして位置づけられた。この方向性の下に、具体的な取り組みとして「公的資金における研究データの管理・利活用に関する基本的な考え方」がまとめられ、研究開発を行う機関の責務の一つとしてデータポリシーの作成が位置づけられている。本講演では、研究デジタル・トランスフォーメーション（研究DX）の推進に向けた政府の政策の動向を解説しつつ、その中での研究データポリシーの役割・位置づけや現状について紹介する。



三宅 隆悟

東京大学大学院情報理工学系研究科修士課程修了（情報理工学修士）。2003年文部科学省入省。原子力、海洋や競争的資金、研究開発基盤等の政策立案に従事。内閣府、内閣官房、復興庁への出向のほか、理化学研究所計算科学研究推進室ではスーパーコンピュータ富岳の設置に向けた企画調整業務を担当。2020年4月より現職。

研究のデジタル・トランスフォーメーション （研究DX）の考え方

私からは、研究データポリシーの作成に当たって一番関連の深い研究DXの施策動向を、海外の動向等も踏まえながら紹介させていただきます。研究DXにもさまざまな議論があり、いろいろな定義がありますが、文部科学省としては以下のように定義して進めています（図1）。

新型コロナウイルス感染症（COVID-19）を契機として、社会におけるデジタル変革やDX化が進みました。研究分野においても、例えば研究機器のスマート化・リモート化、スーパーコンピュータ・ネットワーク等の情報インフラのデジタル化が進んでいます。一方、世界を見ると、やはり知の共有を目指した研究成果のオープン化（オープンサイエンス）が目指されて

いる状況があります。このような状況を踏まえると、研究の在り方そのものについても大きな変革が起きているのではないかと考えているところです。

オープンサイエンスやデータ駆動型研究を推進していくことによって、付加価値の高い研究成果を創出し

研究のデジタル・トランスフォーメーション（研究DX）の考え方

- 新型コロナウイルス感染症を契機として、世界的に社会のデジタル変革（DX）が一気に進展。
- 研究分野においても、研究設備・機器のスマート化・リモート化、サーバ・ネットワーク等情報インフラ利用拡大といったデジタル化が推進。
- 世界的な知の共有を目指した研究成果のオープン化（オープンサイエンス）が進行。
- 研究の在り方そのものに大きな変革期が到来。従来のデジタル化との大きな違いは「研究データ共有・利活用」。様々な研究データが繋がり、AI・データ駆動型研究等に利活用されることで、社会や経済、科学を大きく動かす可能性。

デジタル技術 × 研究データ ⇒ 価値創造
「研究DX」

研究計画・評価分科会（第79回）配布資料（令和4年1月26日）

(図1)

てわが国のプレゼンス向上を目指していくためには、研究データの管理・共有・利活用が重要です。デジタル技術と研究データを合わせることによる価値創造を目指して進めていく必要があると考えています。

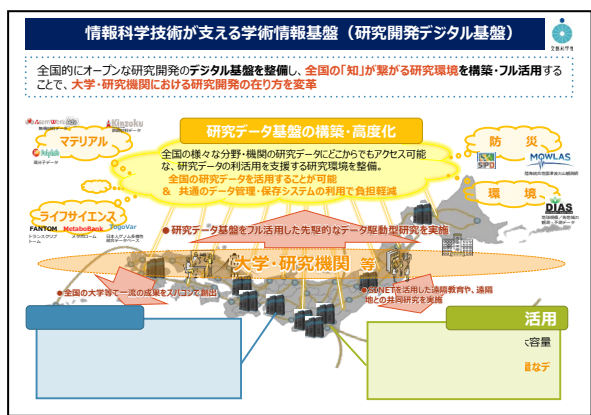
このような状況を踏まえて、研究データの共有・利活用によってデータ駆動型研究による研究プロセスの変革およびイノベータティブな成果の創出を期待しているところです。

背景として、研究データや論文の提供が論文掲載の要件になったり、世界的な出版社や IT 企業がビジネスの対象として興味・関心を高めていることもあり、研究データに関する取り組みは非常に活発化している状況です。

これを踏まえて、そもそも研究データというものが戦略的資源であって、その共有が論文執筆と同等以上の価値を持つという基本的な理念を共有していく必要があります。そのためには、国全体として研究データの共有・利活用ができる基盤・体制を整備し、その中で研究データの共有を促進してデータ駆動型研究を推進するという流れをつくっていく必要があると考えています。これらを早期に起こして動かしていくことで、国際競争力・協調力の向上を図ることを目指すわけです。

情報科学技術が支える学術情報基盤（研究開発デジタル基盤）

もう少し広い観点で言うと、情報科学技術全体で学



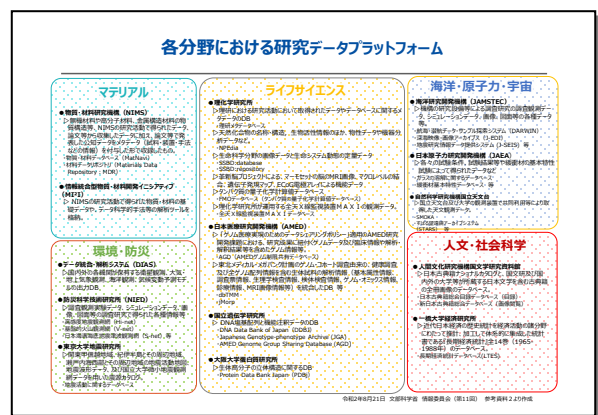
(図 2)

術情報基盤をしっかりと支えていく必要があると考えています (図 2)。研究データ自体を管理する研究データ基盤の構築・高度化という観点もありますし、「富岳」を代表とする国内のスーパーコンピュータをしっかりとつないで計算のリソースを用意する。また、つなぐためにはネットワークが必要なので、学術情報ネットワーク SINET を整備して、これらを一体的に活用して学術基盤を整備し、例えばマテリアル、ライフサイエンス、環境・防災、その他の領域も含めて、研究データの管理・利活用や DX を進めていくということに取り組んでいます。図 3 は、これに関連する研究データプラットフォームの一例です。さまざまな分野において、さまざまなデータプラットフォームが存在しています。こういうものをつなげて研究を加速していくことを目指しています。

オープンサイエンスに係る世界と日本の動向

論文のオープンアクセスと研究データのオープン化を含む概念であるオープンサイエンスは、2013 年の G8 による「オープンデータ憲章」の合意を発端に世界的な潮流となっています。また、それ以外のさまざまな国際機関においても議論が行われ、FAIR 原則に基づくデータの共有や、人材育成・供給などの国際的な対応が求められているところです。

ただ、もちろんオープンといっても全てを公開すればいいのかというと、当初はそういう議論も若干あったと思いますが、そういうわけではありません。例え



(図 3)

ば個人情報や秘密情報、特許のような権利的な話や、わが国にとって重要な情報を含むものなどに関する公開・非公開については、しっかりと戦略的に行う必要があります。

改めて海外の動向を見ると、オープンアクセスの急速な普及に伴って、例えば欧州では Plan S や Horizon2020 のような取り組みでオープンアクセスの義務化の流れが起こっていたり、さまざまな分野で論文のエビデンスとなる研究データについて公開が投稿条件になったりしています。それに合わせて海外大手出版社が論文出版にとどまらず、研究データを含むさまざまな情報を支えるプラットフォームに拡大している状況です。

具体的な議論を紹介します。2021年のG7におけるオープンサイエンスの会議の後、G7研究協約（Research Compact）の中で「データ及びツールを可能な限りオープンかつ迅速に共有する際の障壁を取り除くとともに、オープンサイエンスを促進し、オープンで安全かつ透明性のある形で市民への科学の普及を拡大させ、技術関連のリスクの最小化に努める」ということが掲げられました。こういうものが首脳レベルでの議論の中で出てきたということです。

それを時系列で整理したものが図4です。もちろんこれに至るまでには、2016年のG7茨城・つくば科学技術大臣会合をはじめ、さまざまな議論が行われています。当然、日本においても内閣府や、文部科学省では科学技術・学術審議会、また日本学術会議等々で議

オープンサイエンス/研究データに係る国内外の動向

	2015	2016	2017	2018	2019	2020	2021
G7 日、米、英、仏、独、伊		「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。	「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。				「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。
OECD OECD加盟国		「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。					「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。
ユネスコ ユネスコ加盟国		「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。					「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。
G20 G20加盟国		「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。					「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。
日本		「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。					「データは研究の中心である」として、オープンサイエンスの重要性を強調し、研究データの共有を促進する。

※関係動向等を整理して表記

(図4)

論が行われて、今の政策の状況につながっているという事です。

科学技術・イノベーション基本計画と、基本的な考え方

ここからは政策の中でどのような位置付けがなされているかを紹介させていただきます。図5は、第6期科学技術・イノベーション基本計画の概要です。この中にさまざまな要素が含まれていますが、「新たな研究システムの構築（オープンサイエンスとデータ駆動型研究等の推進）」という章が設けられました。基本計画上も、この取り組みに注力していくということが表されていると考えます。ここでは、研究データの管理・利活用、スマートラボ・AI等を活用した研究の加速や機器の共用、研究DXが開拓する新たな研究コミュニティや環境の醸成をしっかりと進めていこうということが明確にされています。取り組みの柱としてはさまざまな観点から行うこととなっていますが、特に本日の議論に関係するのが、「公的資金により得られた研究データの機関における管理・利活用を図るため、大学、大学共同利用機関法人、国立研究開発法人等の研究開発を行う機関は、データポリシーの策定を行うとともに、機関リポジトリへの研究データの取載を進める。あわせて、研究データ基盤システム上で検索可能とするため、研究データへのメタデータの付与を進める」という部分です。こういった取り組みを基本計画上では求めています。

科学技術・イノベーション基本計画(概要)

現状認識

- 国内外的動向を踏まえ、科学技術・イノベーションを推進する必要性を認識し、国際社会との連携を強化する。
- デジタル技術の進展により、研究データの蓄積・共有・活用が促進され、新たな研究システムの構築が可能となる。
- 科学技術・イノベーションの競争力向上を図るため、研究データの管理・利活用を推進する。

我が国が目指す社会(Society 5.0)

「超スマート社会」の実現を目指す。科学技術・イノベーションを推進し、社会課題の解決に貢献する。

Society 5.0の実現に必要なもの

- 科学技術・イノベーションの推進
- 人材の育成
- 社会課題の解決

Society 5.0の実現に向けた科学技術・イノベーション政策

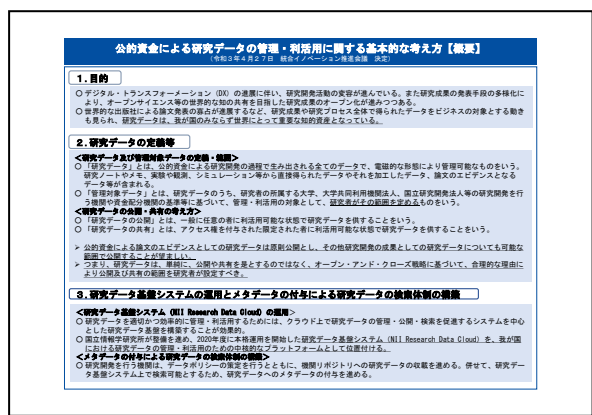
「超スマート社会」の実現を目指す。科学技術・イノベーションを推進し、社会課題の解決に貢献する。

(図5)

それを具体的に進めるために、2021年4月に、閣僚級会議である統合イノベーション戦略推進会議で「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」が決定されています。この中では、研究データ基盤システムを整備し、これを中核的なプラットフォームとして、研究データの管理・利活用にに向けた取り組みを推進していくということが書かれています。さまざまなデータプラットフォームとの連携を推進し、また大学・研究機関等においては研究データにメタデータを付与して収載していく、資金配分機関においても公募型の研究資金の中にその考え方を入れているということが定められています。

「基本的な考え方」において目指す主要な数値目標としては、機関リポジトリを有する全ての大学等において、2025年度までにデータポリシーの策定率が100%になること、また、公募型の研究資金の新規公募分について、2023年度までにデータマネジメントプラン(DMP)およびそのメタデータの付与を行う仕組みの導入率が100%になることが掲げられて、今、進められている状況です。

もう少し詳細をご紹介します(図6)。「基本的な考え方」は、最初に目的をしっかりと書いており、そもそもなぜこういうことをするのかという理念を表現しています。「1.目的」の部分では、研究データはわが国のみならず世界にとって重要な知的資産となっているという考え方を基に、さまざまな取り組みを進めていきたいと思いますというのを書いています。



(図6)

「2.研究データの定義等」では、研究データや管理対象データの定義をしています。ここで大事なのは、研究データそのものという、研究開発の過程で生み出された全てのデータを指しますが、当然それを全部保管したり管理したりするには手間がかかり、無駄も多いです。そのため、管理対象データを決めて、その範囲でしっかりと管理していこうというのがここでの考え方です。

その中で、研究データの公開・共有の考え方として、「公開」とは、一般の方にも見ることができるような共有化の形で研究データを供すること、「共有」とは、アクセス権を付与された限定された人に研究データを供することとしています。ここでは公的資金の話による研究データのことも整理していますので、公的資金による論文のエビデンスとしての研究データは原則公開として、その他についても可能な範囲で公開するのが望ましいとしています。やはり研究データは、単純に公開や共有をすることが目的ではなく、合理的な理由により公開・共有範囲を研究者が設定して、しっかり管理しようということがここで整理されています。当然、これをいったからといってできるわけではありませぬので、さまざまな取り組みを合わせて進めていくことを考えています。

「3.研究データ基盤システムの運用とメタデータの付与による研究データの検索体制の構築」には、NII Research Data Cloudを中核的なプラットフォームとして位置付け、しっかりと運用していく、開発していくことを書いています。

また、各ステークホルダーにおける責務、すべきことを書いているのが図7です。「6.研究者の責務」では、研究者が管理対象データの決定、メタデータの付与、DMPの作成を行い、保存し、オープン・アンド・クローズ戦略に基づいて公開・共有することを求めています。

それを行うための機関としての責務を「4.研究開発を行う機関の責務」にまとめています。データポリシーの策定や、研究データへのメタデータの付与の推進、

それを支援する体制の整備・評価などです。

また、公募型の研究資金における資金配分機関 (FA) の責務については 5 に記載しており、DMP やメタデータの項目の決定、これらの取り組みの評価体系の導入、e-Rad への登録を求めています。

オープンサイエンスの話を中心に説明しましたが、7にあるように、これらの取り組みは研究公正の観点からも非常に重要です。近年、学术论文の査読に当たって、研究公正の観点から研究データの提供やリポジトリでの保存等が求められる場合があります。当然、研究公正の確保は機関や研究者にとって非常に大事なことです。これらをしっかりと説明できるようにするためにも、研究データはデータポリシーに従って適切に保存・管理される必要があります。

研究データ基盤システム (NII Research Data Cloud)

これらに関する施策をご紹介します (図 8)。研究データ基盤システム (NII Research Data Cloud) は、データ管理基盤である GakuNin RDM、検索基盤である CiNii Research、公開基盤である JAIRO Cloud の三つの基盤から構成されており、オープン・アンド・クローズ戦略に対応した研究データ基盤として開発されています。当然ながら、今後管理する上での負担軽減も進めていく必要があるため、例えばデータマネジメントプランに基づくデータ管理支援や大学等研究機関の情報管理といった負担軽減や、研究活動の支援がしや

すくなるような改良・高度化を文部科学省としてもしっかりと支援していきたいと考えています。

研究データポリシー策定のためのガイドライン整備状況について

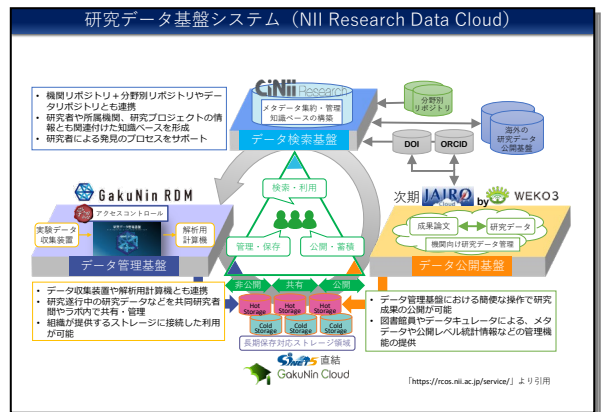
二つのガイドラインをご紹介します。元々のデータポリシーの策定については、2018 年の「統合イノベーション戦略」において求められた関係で研究開発法人が先行して動いています。そして 2018 年に「国立研究開発法人におけるデータポリシー策定のためのガイドライン」が定められて、取り組むべき事項や定めるべき項目が列挙されました。

大学関係については、領域が広く、また自主性の問題もあり、国におけるガイドライン策定はなかなか進んでいませんでしたが、2021 年に大学 ICT 推進協議会 (AXIES) で「大学における研究データポリシー策定のためのガイドライン」を策定していただきました。ここでは、研究データポリシー策定に当たっての留意点等々が整理されています。

データポリシーとデータマネジメントプラン (DMP)

研究開発法人につきましては、全法人についてデータポリシーの策定を目指すということが以前から行われており、文部科学省所管の 7 法人は全て策定済みです。大学の方で先行して策定された事例も把握しています。これらの整備状況も参考にしながら、ぜひ各機

(図 7)



(図 8)

関においてマネジメントプランの検討を進めていただければと思っています。

また、資金配分機関、競争的資金の関係でも順次検討を進めています。図9は文部科学省における研究データマネジメントプランに関する検討状況です。文部科学省が直接公募を行う各競争的資金制度においても、研究データマネジメントプランに関する記載の追記を検討しており、既に追記された公募も先行して行われています。こちらにつきましては、当室と競争的資金の関係を行っている文部科学省の関係部署とも具体的な書きぶりについて相談して、2023年以降は、基本的に新規の公募についてはこのような記載が入る形で準備を進めているところです。

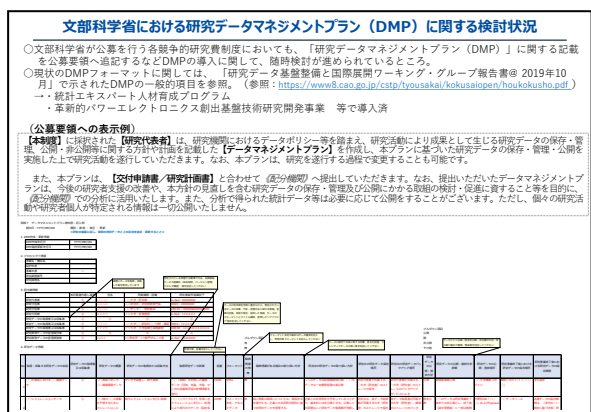
また、資金配分機関でも、先行して動いている例があります。科学技術振興機構（JST）であれば、さきがけ、CREST等で2017年度より実施方針について記載があります。日本医療研究開発機構（AMED）、新エネルギー・産業技術総合開発機構（NEDO）においても2018年よりこの記載がされています。日本学術振興会（JSPS）の科学研究費補助金（科研費）の関係では、2021年度の学術変革領域公募において明示しています。その他の内容についても現在検討中と聞いています。

研究データ公開のポジティブな影響

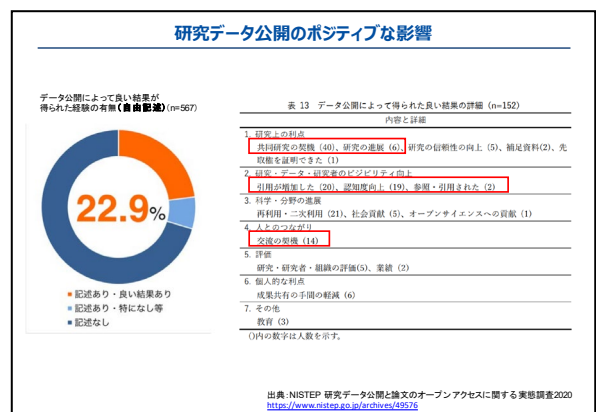
最後に、研究データ公開に関してのポジティブな影響を一つご紹介します。図10は、科学技術・学術政

策研究所（NISTEP）が実施した日本の研究者によるデータ・論文の公開状況への認識を明らかにするための調査（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 2020）のうち、研究データを公開することによって得られるポジティブな影響に関する調査結果を抜粋したものです。今回モデレーターを務めていただいている林先生、池内先生を中心に行った調査と聞いていますが、簡単に紹介させていただきます。左の円グラフを見ると、研究データを公開した経験がある方の中で、公開によって良い影響があったと回答した方が23%いたということです。元々こちらの質問は自由記述ですので、研究データのメリットを確かに感じている研究者の方がいるということだと思えます。右側の表は、研究データの公開の良い影響として回答された内容の詳細です。例えば、共同研究や交流の契機になった、引用の増加や認知度の向上につながったという回答があり、研究データの公開は研究活動にとって確かに良い影響があると感じている先生方もいるということだと思っています。

ここまで、いろいろな内容をご紹介してきましたが、研究データの管理・利活用に向けてさまざまな施策を打っています。文部科学省としても、その管理・利活用がしっかりとできるような形で基本整備を進めたいと思います。各機関、各研究においてこのようなことに取り組んでいただき、さまざまな形で研究データの管理・利活用を進めていただければと思っておりますので、ご協力のほどよろしくお願いいたします。



(図9)



(図10)

◆
◆

●**矢吹** 寄せられた質問を読み上げさせていただきます。「諸外国と比して日本はデータ利活用の面でも既に出遅れているのではないかと漠然と懸念します。日本の学术界が魅力的なものになるためには、特にどのような政策あるいは研究データポリシーが重要とお考えでしょうか。お立場からの回答ではなく、個人的見解でも構いません」。いかがでしょうか。

●**三宅** こちらの研究 DX も含めて、さまざまな取り組みを先行して進められている例はあるかと思います。諸外国においても、やはり具体的な取り組みをしていくことは非常にハードルが高く、研究領域によって、進んでいるところもあれば、なかなか進んでいないところもあるという状況を聞いています。全てを一括して動かしていくのは、このような研究の在り方の変容のようなことを含む話ではなかなか難しいと考えておりますので、まず、できるところから取り組んでいく。文部科学省としても基盤整備をしっかり行い、使い勝手のいい形を準備していくということもありますし、研究の中ではデータ駆動型研究に着目していただく。全員が全員やらなくてはいけないということはありませんが、着目した先生方がそれぞれその研究を進めていき、必要な整備を行っていく。もちろん時間が限られているので、ゆっくり取り組めばいいというものではありませんが、その繰り返しの中で一步一步、着実に実績を積み重ねて広めていく取り組みが非常に重要ではないかと考えています。

SPARC Japan セミナー2021

「研究データポリシーが目指すものとは」

電気通信大学が目指す共創進化スマート社会と その Science2.0 への展開

田野 俊一

(電気通信大学長)

講演要旨



本学は『人間知・機械知・自然知の融合により新たな価値（進化知）を創造し様々な課題を自律的に解決しながら発展し続ける「共創進化機能」を内包した未来社会、すなわち「共創進化スマート社会」の実現を推進している。

キーワードは「3つの要素からなる共創進化機能」にある。即ち、①地球上の全てのデータ（センサ（自然、社会）、情報システムが作るデータなど）、機能（機械・機器などの物理的な制御機能、情報システムの機能など）をインターネット経由でアクセス可能とし、②莫大なデータ、機能に対して、AIの機能を活用しつつ、人間知、機械知が連携して分析することにより、新たな知見（サービス、機器などの制御方法、情報提供方法、診断方法など）を発見し、③発見した知見を安定性（社会、システム、制御などの観点）と監視可能性、制御可能性を検証後、社会に埋め込み、社会を進化させる。この「①⇒②⇒③⇒①⇒…」が文字通りリアルタイムに繰り返され社会が自律的に進化し続ける。

この仕組みは科学的発見の形態まで変える。これまでは人間が努力し科学的発見を行ってきたが、全世界のデータや機能をネットワークを介して自由自在に活用すること、および機械知（AI）との共創により、科学的発見をも加速することが期待される。

田野 俊一



国立大学法人電気通信大学学長。博士（工学）。1983年東京工業大学大学院総合理工学研究科システム科学専攻修士課程修了。同年（株）日立製作所システム開発研究所入社。1990-91年カーネギーメロン大学客員研究員。1991-95年国際ファジィ工学研究所。1996年電気通信大学大学院情報システム学研究科助教授。2000-01年マサチューセッツ工科大学客員科学者。2002年電気通信大学 教授。人工知能、あいまい理論、認知モデル、知的HCIの研究に従事。

はじめに

先ほど研究データポリシーという言葉が随分出ていましたが、私はそういう真面目な話ではなく、実際の応用面で電気通信大学が進めている試みと、それが Science2.0 に関係するというお話をしたいと思います。具体的には、本学が目指す共創進化スマート社会とはどのようなものかということと、実際に行っていること、実は Science2.0 に共創進化スマート社会を目指した試みと Science2.0 への発展の二つが繋がっているということ、そして、実際にはなかなか Science2.0 に

は行きませんので、それを解決するためにはこんな課題があるということをお話しします。きっと Science2.0 を本格的に動かすためにはこの課題が関係するのだろうという気持ちでお話ししようと思っています。

電気通信大学が目指す共創進化スマート社会

最初に、共創進化スマート社会についてです。2021年7月に電気通信大学は新しいビジョン（UEC ビジョン ~beyond 2020~）を決めました。ビジョンでは、

本学は Society5.0 を「共創進化スマート社会」と考え、その実現のためにどのような新しい機能（共創進化機能）が必要かを述べています（図 1）。これはバズワードを並べたように見えますが、実際はもう少し細かい記述があります。共創進化スマート社会とは、AI、ネットワーク、ロボット、量子技術などによって自律的に進化する社会です。日本政府が目指す Society5.0 は幅が広くてよく分かりませんが、本学ではこのような技術で自律的に進化する社会を Society5.0 と捉えて、その教育・研究・実現をすること、自らもスマート大学になることをミッションとして挙げています。

実際には、次の三つの機能で社会自身が自律的にイノベーションを起こし発展することを計画しています（図 2）。

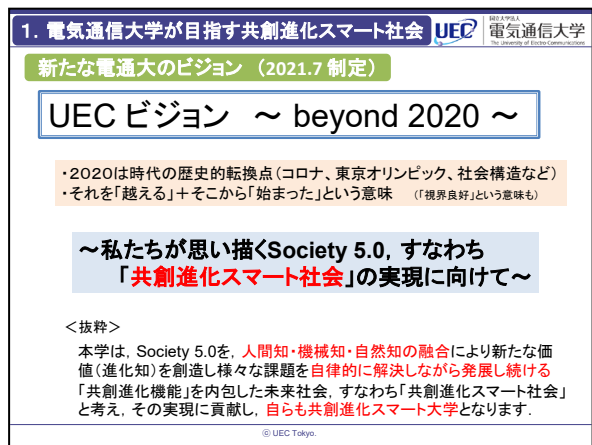
一つ目は、データと機能を地球規模でネットワーク連携させることです。「センサデータ」と書いてありますが、社会的なデータ、実験のデータ、人間が動く実際の自然の社会のデータなど、全てのデータです。機能とは、例えば少し危ないのですが、電車を止める、動かすといったものです。そういうデータと機能を地球規模でネットワーク連携する。データだけではなく機能までつながってしまうということです。

二つ目は、今でもそうなのですが、データが集まり過ぎていて人は見られず、捨てているだけです。たくさんのデータとたくさんの機能を機械知に投げて、どんな新しいルールが見つかったか、どんな

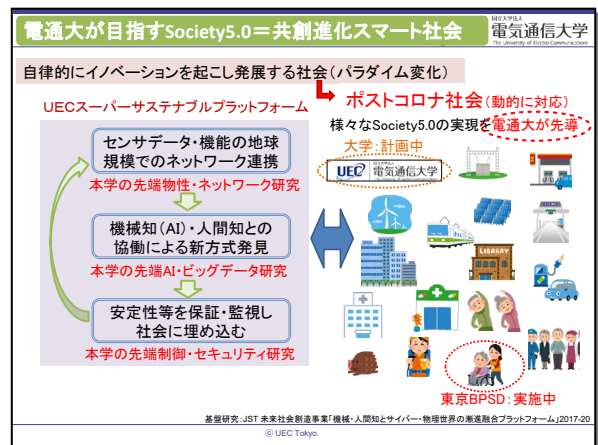
面白い観点の知見が得られたかを、まず機械知と人間知が話し合って新しい方法を見つけるということです。

三つ目として、それが見つかったら、そのまま社会に出てしまうと社会が爆発するかもしれませんので、安定性や可制御性、可観測性を検証して社会に埋め込むということです。埋め込むと、また状態が変わって、データが変わって、機能が変わって、新しいものを見つけてというように、どんどん新しくなる。そのようなものを私たちは共創進化スマート社会と呼んでいます。本学の研究テーマである「先端物性・ネットワーク」、「先端 AI・ビッグデータ」、「先端制御・セキュリティ」に取り組もうということです。対象は、例えば電力でもいいですし、公共の電車でもいいですし、おじいちゃん・おばあちゃんでもいいですし、どこでも同じような仕組みでリアルタイムに進化するだろうと思っています。

この社会的なインパクトは、自律的に進化し続ける社会が出現するという事です（図 3）。今までは、設計 1 年、実装 1 年、運用 10 年などというって、10 年ぐらいつつしか良くならない社会でした。しかし、共創進化スマート社会であれば、進化の源は万人と機械知になり、進化がリアルタイムに進んでいくだろうと思っています。これまでは特定の企業がデータを握って社会進化を担っていましたが、そうすると特定の企業の立場での最適化にとどまりますし、大企業中心になります。それが、多数のベンチャーが社会を担うようになると、どんなことが起こるのでしょうか。全ての



(図 1)



(図 2)

データが共有されていくと、公開されたものをうまく使って、多数のベンチャーが担う新しい社会ができてくるとしています。実現するためには随分お金がかかるのですが、私たちはこんなことを考えて電気通信大学の目標としています。

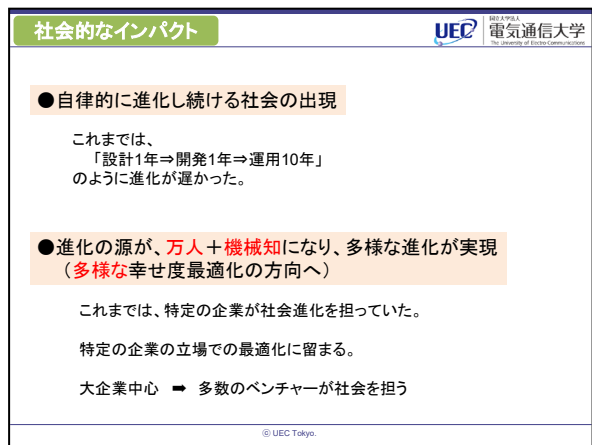
共創進化スマート社会を目指した試み

次に、どんなことをしようとしていたかと、今進めていることをお話しします。

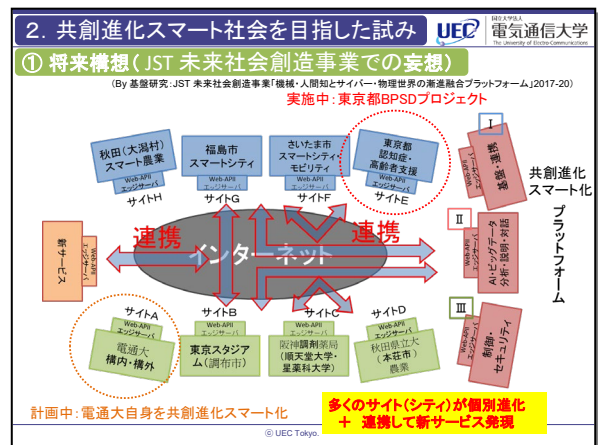
最初は科学技術振興機構（JST）で行っていた未来社会創造事業「機械・人間知とサイバー・物理世界の漸進融合プラットフォーム（2017-2019 年度）」です（図 4）。秋田県大潟村のスマート農業や、福島市のスマートシティ、さいたま市のスマートシティ・モビリティ、阪神調剤薬局、電気通信大学の中でのデータを使って、データと機能を連携する、機械知（AI）と人間知で分析する、社会に埋め込むという先ほどの三つの機能をうまく使うことで、各サイトの事業が自律的にどんどん進化していく、特定の電車が進化していくということです。この事業で目指したのは、あるサイトのデータと連携させることによって、別のサイトで新しいサービスが生まれることです。個別のシステムを活性化させるのではなく、全然違うところのデータを使って自分のところが進化し、全く新しいサービスができるということを目指して進めていました。しかし、チームはつくったのですが、話だけで終わっています。

そこで、一部分だけ進めようということで、電気通信大学と東京都で東京都 BPSD プロジェクト（AI と IoT により認知症高齢者問題を多面的に解決する東京アプローチの確立）として実装しています。BPSD とは、急に怒り出す、急に徘徊するといった認知症患者の行動・心理症状のことです。これがあると、お父さんが急に怒り出した、変になってしまったということで家族が疲弊しますし、罵声を浴びせられてもう嫌だといって介護士がやめていきます。従って、BPSD の予測と対処はとても重要な問題です。

BPSD に対して、普通は医学的知識を使って「こんなときに怒り始める」と対処するのですが、東京都では違うアプローチを取っています。患者の介護施設や病院、自宅など、いろいろなところにセンサー類を設置し、「今お父さんが怒った」、「今こういう体操をしたら落ち着いた」というデータを全て集めて AI で関連付けます。そうすると、例えば心臓がばくばくする徐脈が 3 回連続でどれぐらいの範囲で起こると 50 分後に怒り始めるなど、医学的知識では分からなかったことをどんどん発見し始めます。囲碁や将棋の世界で AI は人間と全然違う戦略をとっていますが、同様に、医学的知識ではなく、とにかくデータを分析して新しい知識を導き出します。得られた仮説を医療・介護専門家が分析し、家庭内のエッジシステムに情報を送り、例えば「30 分後に認知症の障害が出て家を出てしまうよ。注意してね」、「怒りそうだよ。この音楽を聴くと落ち着いたことがあるよ」といった解決策を出しま



(図 3)



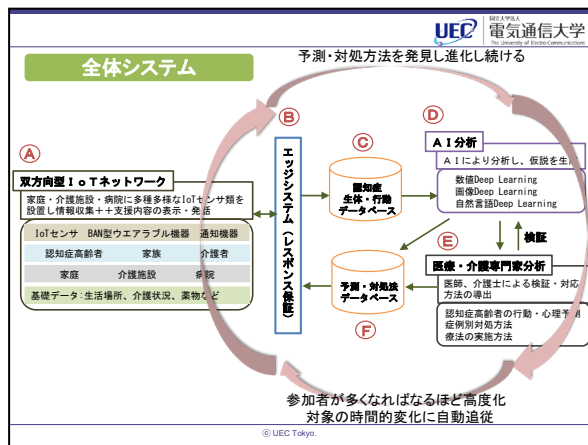
(図 4)

す。

具体的な内容は図5のとおりです。たくさんのデータをどんどん蓄えていき、AI分析をしていろいろな仮説をつくる。そして、つくった仮説をそのまま家庭に送るのではなく、医療と介護の2人の専門家が検証し、OKが出たものを対処法のデータベースに入れると、それが延々と進化していく。そういったことを目指しています。これは事業期間が3年間で、東京都と大学が連携し、民間企業の協力のもと、一緒に取り組んでいます。今は2年目で、データが集まってAI分析を始めているところです。このような進化型の取り組みを実際に進めています。このBPSDプロジェクトでは、参加者はデータを結構提供してくれます。(分析結果が)自分に跳ね返ってきますので、データは戻ってきます。しかし、他の事業では、なかなか提供いただけません。自分の提供したデータが他の人に使われるとなると難しい。データの利活用はとても難しい問題なのです。

東京BPSDプロジェクトを本当は社会で実装したいのですが、社会の縮図かつ協力してくれるところはどこかと考えたときに、電気通信大学だということで、まずは本学で同じような取り組みを進めているところです。

本学には、研究者情報システム、教務システム、財務システム、設備管理システムといったデータベースがあります。これを起点に、学内のIoTやAIブリッジによって、図書館や教室にある人感センサー、CO₂



(図5)

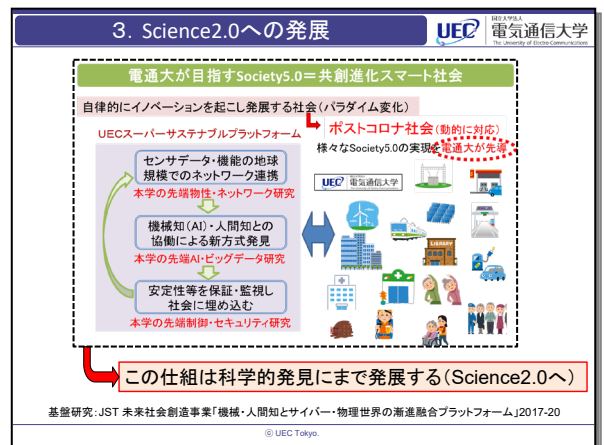
センサー、照度センサー、発電量センサーなど、いろいろなセンサーのデータを分析して、知見を見つけて埋め込むというシステムを回しています。つまり、電気通信大学を一つの社会とみなすということです。本学はデータの所有者ですし、教職員はICT系でこういったものへの理解が深いので、今、幾つかの建物でこの取り組みを進めていますし、いろいろなデータベースも将来的には連携したいと思っています。

そうすると、例えば新型コロナウイルス感染症(COVID-19)対策が見つかるだけではなく、感染者を検出したらドアをロックして強制換気が強くなるということが自動的に動き始めたり、エネルギー効率化が自動的にアップデートされたり、大学での効果的な教育方法を見つけてカリキュラムをどんどん変えていたり、図書館の開館時間も変わるかもしれないということを進めています。

目指す共創進化スマート社会に向けてたくさんのサイトで取り組みをしたいのですが、なかなかデータが出てきませんので、今は特定の問題でデータをシェアすることと、電気通信大学という一つの小さな社会で取り組みを進めているところです。

Science2.0 への発展

このように三つの機能をぐるぐる回して社会を変えようとする仕組みは、実は科学的発見まで発展し、Science2.0につながるという話を JST 未来社会創造事業のときからしていました (図6)。これまでの実験



(図6)

は、データの収集、分析、発見を科学者個人の努力でやってきました。ですから、莫大なデータや機能を科学者個人が一生懸命見て、これとこれは関係しているということを見つけてきたのですが、その隣に助言者として機械知（AI）を置くと、面白い発見の手助けをしてきて、きっと「センサデータ・機能の地球規模でのネットワーク連携」と「機械知（AI）・人間知との協働による新方式発見」の部分でサイエンスが加速化するだろうと思っています。

そうなれば、どのような科学的発見が起こるかという、例えば一つ目は浮力です。お風呂の水があふれたのを見たアルキメデスが、「そうだ、浮力とは出ていった水の量だ、EUREKA!」と叫んだ自然現象です。

二つ目はフィボナッチ数列です。 $F_0=0$ 、 $F_1=1$ 、 $F_{n+2}=F_n+F_{n+1}$ ($n \geq 0$) の形で、前二つの数字を足すという単純な数列ですが、実は花の花卉の数など、非常に多種多様なところで出てきています。こういうものを見つけて、例えば AI に「このようなフィボナッチ数列で説明できるところを探さない」と指示すると、全世界のデータが見えて、実はアリもそうだった、カメもそうだったということを見つけてくれます。

三つ目は複雑系です。鳥が並んで飛ぶことは、実は数個の行動ルールで記述できることが発見されてから複雑系の科学が進みましたので、社会的なデータであっても、そこからいろいろなデータの見方や規則性を AI が発見してくれると、科学的発見に寄与するのではないかと考えています。

人間と機械知と自然知の連携例としては、人間が仮説を得て、それと同種のデータが世界中のどこかにないか機械知に探させるとか、複雑なものが数個のルールで表せるという面白い仮説を機械知が発見して人間に提案するということがあります。このような連携が科学的発見を加速させるのではないかと考えています。

解決すべき課題とアプローチ

そういうことを考えているのですが、実はなかなか

うまくいきません。そこで、解決すべき課題とアプローチについてお話しします。これは、きっと皆さんがしようとしているような、オープンデータで科学的なものを法則化するときに使えるのではないかと思います。

課題は三つあります。一つ目は、多くの基盤技術が必要となることです。個別 ICT について、連携、AI、安全性をいろいろな学問分野で研究しなければなりません。理工系だけでは駄目で、アート系からビジネス系、政策系まで含めて研究しなければならないと思います。また、データの提出を促すような政策も重要です。（政策関係者・研究データ流通関係者などの）皆さんの例で言うと、公的資金による研究で出たデータは公開するという政策です。政策でやらないと、なかなかデータは集まらないと思います。

二つ目は、「社会」と呼べるような実験場が必要なことです。かなり大きなデータがないとつまらないですし、多くの多様な人がデータを集めてくれないとうまく回りません。どう集めるかが問題です。

三つ目は、社会構造の大転換となり、変化への抵抗が大きいこと、つまり反対する人が絶対に出てくることです。サイエンスのやり方が変わるとそれに反対する人が出てきます。特に、既得権益を守りたいところはなかなかデータを公開してくれません。データを公開すると他の人のためになるかもしれませんが、公開しないことには真のイノベーションが生まれません。従って、既得権益を守りたい人たちをどうにかして、たくさんのデータを集めないとなりません。

では、どのようなアプローチを取るべきでしょうか（図 7）。一つ目の課題については、多様な研究が必要です。二つ目の課題については、特定の社会に適用した典型的な成功事例が必須です。それがあると、みんながデータを公開してくれるかもしれません。公開すると良いことがあるということを示さないということでは、三つ目の課題については、データ機能のオープン化など、政策の整備が必要です。そしてもう一つ必要なのは、誰でも自由に使えるという民

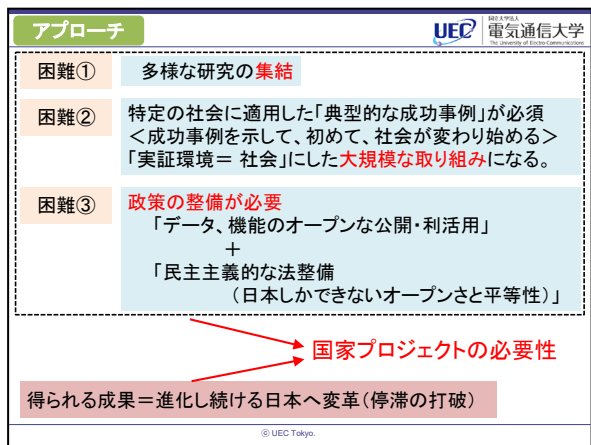
主主義的な法整備です。国が強制させて使う方法ではなく、民主主義的な法整備をしなければならないと思います。共創進化スマート社会の例で話をしていますが、皆さんのオープンサイエンスと同様に、国家プロジェクトで取り組まないと進まないだろうと思います。

政策の必要性

最後に政策の必要性についてです（図 8）。欧米各国の EV 化は政策で進めています。他には、オープン API 化によって銀行のデータが出てきています。EU が銀行のデータと機能をオープンにすることを、（第二次決済サービス指令（PSD2）において）義務付けたのです。これにより、Web API を用いて、データ（お金を幾ら持っているか）と機能（どこからどこまでお金を振り込んだか）がオープンになりました。オープンになったことで、いろいろなベンチャーが発達して新しいフィンテック企業が台頭し始めています。ですから、こういった政策がとても必要だと思います。

●矢吹 質問が一つ来ています。「先生がもし人文社会科学系学部を擁する総合大学の学長だとしたら、本日の講演内容を受けてどのような大学運営をされるでしょうか」。

●田野 例えば文系の人とはこの話をしたいと思って



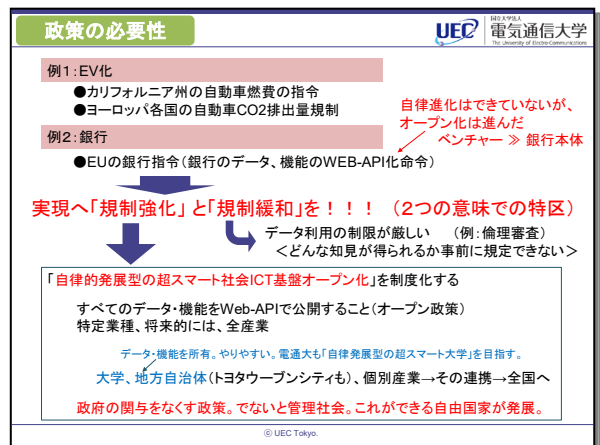
(図 7)

います。社会をどう変えたいかというところは実は文系です。私たちは AI や ICT のことはできるのですが、ビジネス系や政策系、実際に社会がどうなってほしいのかというところはかなりアートに近く、私たちはない部分です。そういうところとの連携を強化しています。

●矢吹 もう一つ質問があります。『スマート大学』というワードにわくわくします。新技術によるスマート化や新サービスは研究・教育面では率先して導入されるものの、事務方にはなかなか恩恵がないイメージがあります。人が持つ習慣、変わることへの心理的抵抗感を超えて、大学構成員が一丸となって共創進化するにはどうしたらいいとお考えでしょうか。

●田野 私が学長になる 2、3 年前は、本学はスマート大学を目指して、全てのデータが集まるとか、データベースを整備しようと言っていました。これは結構反対が多かったため、仕方なく外部コンサルタントに依頼して、今がどれだけ非効率かということで説得しようとしていました。しかし、現在はそのようなことはありません。なんと COVID-19 感染拡大の影響により、全てのデータをつながないとまずい、全て自宅でするようにしないと困る、そのためには仕事の仕方を変えなければならないということを事務職員の方々はほぼ分かってくれました。

従って、COVID-19 を契機にスマート化は非常に加



(図 8)

速すると思います。昔は、デジタル化すると仕事の仕方が変わると反対が増えたのですが、デジタル化しなければ在宅勤務ができず効率が悪いと、みんながやりました。COVID-19 の時代がチャンスだと思っています。

SPARC Japan セミナー2021

「研究データポリシーが目指すものとは」

学術情報インフラが実現する 研究データの管理と循環

大波 純一

(国立情報学研究所 オープンサイエンス基盤研究センター)

講演要旨



研究データポリシー実現のためには、研究データを適切に管理する学術情報インフラの環境が必要である。一方これを実現するため、大学を含む研究機関は新たな仕組みの構築や基準を策定する必要に迫られた。そこで国立情報学研究所では公的な情報基盤を構築し、利用者の研究データ管理の負担軽減となるように対応を開始している。NII Research Data Cloud (NII RDC) と呼ばれるこの基盤は、多様なアプリケーションとしての機能と高い処理能力を有し、国内の機関や研究者の研究データ管理をサポートし、研究活動における知の循環を促進する。このNII RDCを利用した研究データ管理や、研究データの公開・検索を含めたプロセスについて紹介する。

大波 純一



国立情報学研究所オープンサイエンス基盤研究センター特任准教授。博士（理学）。2009年3月、東京工業大学大学院生命理工学研究科博士課程を修了。製薬企業向けのシステムエンジニアとして民間で働いた後、国立研究開発法人科学技術振興機構バイオサイエンスデータベースセンターで研究員として、生命科学系データベースや検索基盤の研究に当たる。2020年4月より現職で学術検索基盤CiNii Researchの研究開発、運営および国際連携に従事。

はじめに

本日は国立情報学研究所（NII）が提供しているインフラと、研究データポリシーや研究データの管理に関する状況についてご説明します。

最初に少し自己紹介をさせていただきます。私は元々、生物学の方で学問を修めており、動物の遺伝子解析のようなことをしていました。卒業後はIT企業で製薬企業向けのシステム開発などを行い、生物系や情報工学の分野ではとても多くのオープンデータを使ってきたことから、オープンサイエンスの重要性を認識しました。その後、公共データベースの現場へ移り、2012年から科学技術振興機構バイオサイエンスデータベースセンターで横断検索などのシステムを開発し

ました。2020年に国立情報学研究所のオープンサイエンス研究センターに移り、現在は新しい検索基盤のCiNii Researchを担当しています。本日は、この基盤の説明からさせていただきます。

学術コンテンツ流通の広がり

元々、学術コンテンツといえば論文や図書、学位論文などが主流だったと思いますが、最近はオープンアクセス推進の流れがあります。国際社会のさまざまな要請もあり、実際にそれをどうしていくかという段階になっています。さらに、研究データを研究データ管理（RDM）の流れで出したり、歴史的・文化的資料についてはデジタルアーカイブで公開したりといった、

幅広い提供の仕方をしていくという世界が学術コンテンツにおいて広がってきています。

このような中で国の政策としては、機関リポジトリを有する全ての大学、大学共同利用機関法人、国立研究開発法人において、2025年までにデータポリシーの策定率が100%になるようにしてくださいということが提言されました。さらに、公募型の研究資金の新規公募分において、2023年度までにデータマネジメントプラン（DMP）およびこれと連動したメタデータの付与を行う仕組みの導入率が100%になるようにしてくださいということもいわれるようになりました。このことは、国の示している第6期科学技術・イノベーション基本計画の方針で指示され、それに従って研究・管理をしていくという動きが必要になってきています。

実際にどのような動きをするかということも国で定めています。例えばデータキュレーター、図書館職員、URA、研究の第一線から退いた方など、研究の周りに位置しているさまざまな人たちの協力を頂き、図書館のデジタル転換などの取り組みについて、2022年度までに方向性を定めていくということです。

そこで、ここからは、「公的資金による研究データの管理・利活用に関する考え方」、研究データ基盤システム NII RDC（NII Research Data Cloud）、およびメタデータ共通項目と呼ばれるデータのテンプレートについてご説明します。

「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」

「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」は、2021年4月に統合イノベーション戦略推進会議にて決定されました。この内容は内閣府の研究デジタルトランスフォーメーションのWebページで公開されていますが、主に図1に示す四つの記述のとおりです。

一つ目は、研究データ基盤システムである NII Research Data Cloud を中核的なプラットフォームと位置

付けて、産学官における幅広い利活用を図るため、メタデータ（データを説明するための情報から構成されるデータ）を検索可能な体制を2023年度までに構築することです。

二つ目は、研究開発を行う機関は、データポリシーを策定し、機関リポジトリへの研究データの収載を進めるということです。

三つ目は、公募型の研究資金の全ての新規公募分について、DMP（研究データの管理計画書）やメタデータを付与する仕組みを2023年度までに導入することです。

四つ目は、研究者は、所属機関のデータポリシーや公募型の研究資金における資金配分機関の基準などに基づき、管理対象データの範囲を定めてメタデータを付与し、研究データ基盤システム上において検索可能となるように登録する。そして、研究者や研究プロジェクトをマネジメントする者は、資金配分機関の求めに応じて、DMPを原則として研究開始前に作成し提出するという事です。

しかし、これらの項目には、研究データのメタデータやDMPの作成、機関リポジトリへの収載、2023年度までに行うなど、さまざまな新しい内容が含まれており、これらを完璧に準備できている機関はまだ少ないと思います。

図2は「公的資金による研究データ管理・利活用に関する基本的な考え方」についてのスライドです。特に注目していただきたいのは、真ん中に赤で示した研

NII RCOS

主な内容

- 研究データ基盤システム（NII Research Data Cloud）を中核的なプラットフォームとして位置付け、産学官における幅広い利活用を図るため、**メタデータ**（データを説明するための情報から構成されるデータ）を検索可能な体制を構築する。（2023年度まで）

- 公募型の研究資金の全ての新規公募分について、**DMP及びメタデータ**を付与する仕組みを導入。（2023年度まで）

- 研究開発を行う機関は、データポリシーを策定し、機関リポジトリへの研究データの収載を進める。（機関リポジトリを有する全ての大学・大学共同利用機関法人・国立研究開発法人においては、2025年までにデータポリシーを策定）

- 研究者は、所属機関のデータポリシーや公募型の研究資金における資金配分機関の基準等に基づき、管理対象データの範囲を定め、**メタデータ**を付与し、研究データ基盤システム上において検索可能となるように登録する。
- 研究者や研究プロジェクトをマネジメントする者は、資金配分機関の求めに応じて、**DMP（研究データの管理計画書）**を原則として研究開始前に作成する。

（図1）

究データ基盤システム（NII Research Data Cloud）です。ここに、大学・研究機関や分野ごとのデータプラットフォームの中から、メタデータやデータそのものを登録していただき、それを検索・利活用できるようにしてアカデミアや産業界に出していくことを推進しようとしています。この結果、先ほど電気通信大学での Society5.0 につながるさまざまなアウトプットの話があったように、さまざまな方向から社会が良くなることを実現していこうという流れになっています。

このデータの流れを実現するために、資金配分機関や NII ではメタデータを収載する際の基準となる共通項目を制定しています（図 3）。こちら内閣府の Web ページで公開している内容ですが、「ムーンショット型研究開発制度におけるメタデータ説明書」でパイロット的にデータ登録している分野において、メタデータのテンプレートを設定しました。これを基準と

して、どのような項目のメタデータを研究データから登録して流通させるかを考えています。

研究データ基盤システム（NII Research Data Cloud）

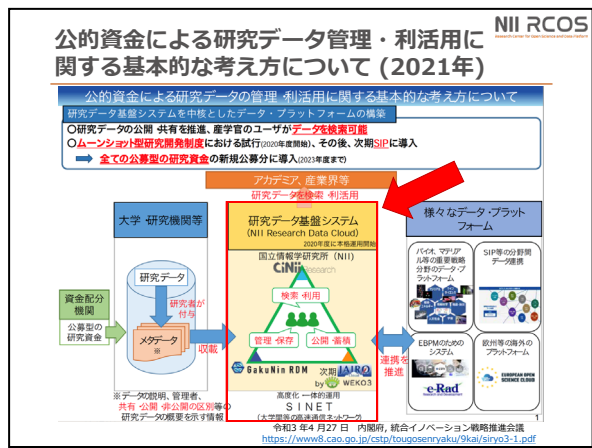
さらに、NII では近年、NII Research Data Cloud（NII RDC）という基盤を提供しています（図 4）。これは大学や研究機関で創出された学術情報を NII のクラウドにまとめていただき、活用・保存していくための場所です。大まかには、データ検索基盤、データ管理基盤、データ公開基盤の三つで成り立っており、それぞれが連動しています。

データ管理基盤では GakuNin RDM と呼ばれるサービスが動いており、さまざまなデータの登録やプロジェクト内のデータ管理ができるようになっています。

データ公開基盤では、JAIRO Cloud というクラウド環境、もしくは WEKO3 と呼ばれるサービスにデータを登録して外に公開することができます。

そして、データ検索基盤では CiNii Research というサービスが動いています。流通しているメタデータをこちらに集約し、利用者は検索して情報を発見することができます。

NII RDC を実際に使うときの流れは図 5 のようになっています。基本的には研究者などが GakuNin RDM に日々の研究データを管理・共有するためのスペースを作って登録します。メタデータや研究データ本体、そして DMP の情報も登録できます。さらに公開基盤

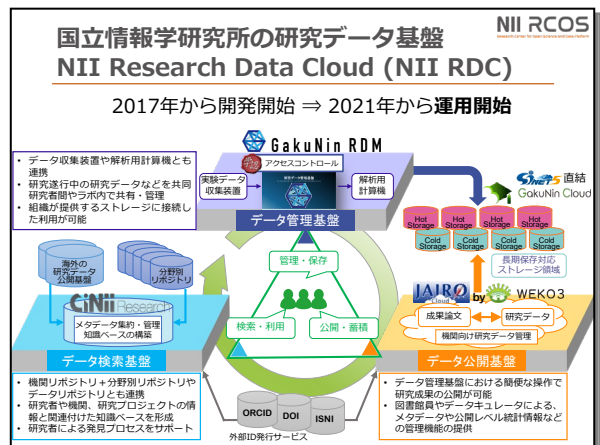


(図 2)

項目	必須/任意	備考
1 資金配分機関情報	必須※	公的資金の研究資金を配分した資金配分機関 (府省含む) の実施機関 公的資金の研究資金以外の場合は記入不要
2 e-Radの課題番号	必須※	公的資金の研究資金の場合は、e-Radに登録した課題番号 公的資金の研究資金以外の場合は記入不要
3 プロジェクト名	必須※	プロジェクトの研究代表者が所属する研究機関の機関名 公的資金の研究資金以外の場合は、研究機関の機関名に記入
4 データNo.	必須	管理対象データを一意に特定するための番号
5 データの名称	必須	研究データ、報告資料、測定結果などの呼称を記載する 公開予定日
6 掲載日・掲載更新日	必須	データの掲載日、掲載更新日
7 データの説明	必須	随時かつ簡潔に内容を記載
8 データの分野	必須	e-Radの研究分野 (主分野) ・e-Radとの連携により、自動入力される予定。
9 データ種別	必須	研究データ基盤システム上では、通常は「データセット」を標準とするが、データの特性に応じて「データファイル」の種別を選択可能
10 掲載データ量	任意	1GB未満、1GB以上10GB未満、10GB以上100GB未満、100GB以上の区分により記載
11 管理対象データの活用・提供方針	必須	無償/有償、ライセンス情報、その他条件 (利用の仕方等) を記載
12 アクセス条件	必須	公開/共有/非共有、若くは公開/公開範囲外から選択
13 データの公開範囲	必須	公開範囲を選択した場合は、公開予定日を記載
14 データの公開範囲	必須	特定のIPアドレス、特定のIPアドレス以外のIPアドレスに限定
15 データ作成者	任意	管理対象データを生み出した研究者の名前
16 データ作成者のDe-Rad研究者番号	任意	管理対象データの作成者のDe-Rad研究者番号
17 データ作成者のDe-Rad研究者番号	必須	De-Rad研究者番号を記載
18 データ管理機関 (De-Rad) 研究者番号	任意	e-Rad研究者番号
19 データ管理機関 (De-Rad) 研究者番号	必須	e-Rad研究者番号
20 データ管理機関 (De-Rad) 研究者番号	必須	e-Rad研究者番号
21 連絡先	任意	データ管理者

※公的資金の研究資金による研究活動の場合

(図 3)



(図 4)

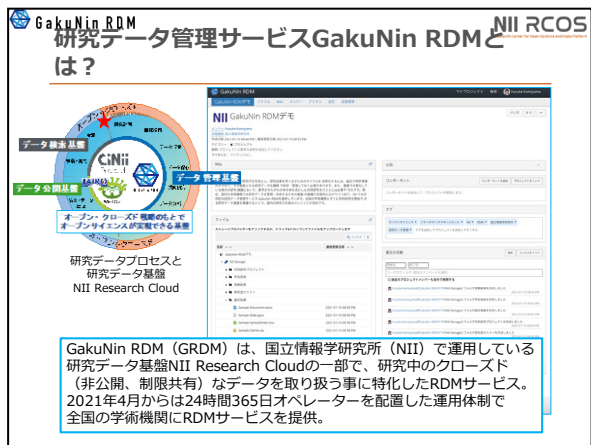
では、研究データを実際に外に出すための設定をします。大学では図書館員などがその処理に当たるとこともあると考えています。最後に、そのようにして公開されたデータを検索基盤の CiNii Research で外に出すと、組織管理者や外部研究者が確認したり、一般市民の方が情報を見たりすることができるようになります。そういった使い方を想定しています。

管理基盤 GakuNin RDM

管理基盤 GakuNin RDM は、登録すると図 6 のような画面になり、プロジェクトの中でデータ共有やデータ流通などをすることができます。まだ構築中の部分もありますが、実際に登録するときはフォームに入力していただきます。さまざまな互換機能も内部に備えています。メタデータの共通テンプレートに入力するのは大変だと思った方もいると思いますが、できるだ



(図5)



(図6)

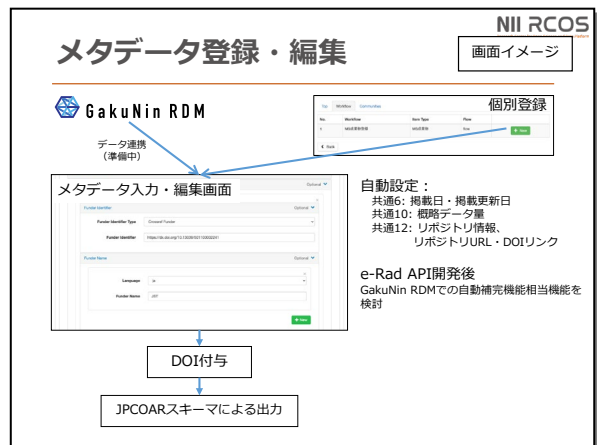
け Web のインターフェイスで入力しやすいように設定しています。日本語、英語の対応や、さまざまなスキーマへのフィッティングなども自動で行うことを検討しているため、お使いいただければと思っています。最終的に登録した機能をレポートニングすることもできる予定です。DMPの登録なども予定しています。

公開基盤 JAIRO Cloud

図 7 は公開基盤 JAIRO Cloud の画面のイメージです。登録されたデータ一つ一つを論文に出すとき、あるいは最終的に世の中に広く見てもらいたいものがあるときはこちらに登録し、DOI を付与して、スキーマなどを入力して見せることができるようになっています。

検索基盤 CiNii Research

検索基盤 CiNii Research は 2021 年 4 月から正式公開された国立情報学研究所の新サービスで、一つの検索画面から多様な学術情報へイージーアクセスできる基盤です (図 8)。3,000 万件以上のデータが入っており、論文だけではなく、さまざまな研究データ、本、研究のアウトプットを検索することができます。元々NIIが提供していたさまざまなサービスもあるのですが、そのまま掲載したわけではありません。例えば、Google で「CiNii Research」と検索すると CiNii Research の現在のトップ画面にアクセスできますが、その中に「研究データ」や「プロジェクト」というタブがあり、必要に応じてフィルターをかけることができるように



(図7)

なっています。また、検索結果の画面では、データソースとしてどこから資金配分がなされたかでフィルターをかけたり、その研究データがどういったプロジェクトでつくられたかを記述したりと、高度な使い方ができるように工夫しています。

このような研究データおよびメタデータの登録については、ムーンショット型研究開発制度の情報をパイロット情報として、内閣府と登録の方法を構築しつつあります。この中では、特にメタデータの登録の推奨案として、研究プロジェクトの管理者や研究者がまずは管理基盤 GakuNin RDM にデータを登録し、そのデータを公開するときは JAIRO Cloud にそのまま連携させて、そのデータが検索基盤の CiNii Research に移り、自動的に外から検索できるようにして公知にしていくというサポートの仕方を構築中です。

サービス統合のお知らせ

最後に1点、お知らせです(図9)。元々NIIの基盤で論文を検索するためのインフラとして提供していた CiNii Articles を多くの方にお使いいただいていたと思いますが、こちらを2022年4月1日から新しい論文検索基盤 CiNii Research に統合する予定です。実際は、サービスのドメインなどが主に変わって、元々の CiNii Articles の URL はそのままリダイレクトされて CiNii Research の中のページに移るので、ユーザー側で大きな対応をしなければいけないということは少ないと考えています。ただ、API や高度な利用をしてい

た方については少し変更点があります。大変お手数ですが、CiNii Research もしくは現在の CiNii Articles のページに、各ページの右上にお知らせリンクが出ていますので、そちらの内容を見ていただければ幸いです。

一例として、CiNii Articles と CiNii Research のページを、ある同じ論文について示してみます(図10)。CiNii Articles では左のようなメタデータが見えていますが、新しい CiNii Research では、メタデータだけでなく、参考文献のリンクや関連プロジェクト、ものによっては関係する研究データなどつながって、そのままシームレスにさまざまなデータを見つけたり使ったりすることができるようになっています。ご利用いただければ幸いです。

まとめ

発表のまとめです(図11)。まず、「公的資金によ

(図9)

(図8)

(図10)

る研究データの管理・利活用に関する基本的な考え方」(2021年)に沿う研究データの管理が必要な状況になってきました。それに対応して、メタデータ共通項目にマッピングされた研究データを、研究データ基盤システム(NII RDC)の三つの基盤である管理基盤、公開基盤、検索基盤を利用して適切に管理し、新たな知の循環に持っていくことができると考えています。最後に、2022年4月にCiNii ArticlesはCiNii Researchに統合します。より発展的に使うことができるものになっていますので、よろしくお願いたします。

●矢吹 二つ質問を頂いています。一つ目は、「GakuNin RDMとJAIRO Cloudの連携では、メタデータだけでなく、研究データ本体の連携が可能な見込みでしょうか」という質問です。いかがでしょうか。

●大波 研究データ本体の連携が可能な見込みです。研究データ本体は、管理基盤、公開基盤まで移って、正確にはメタデータもJAIRO Cloudに移るのですが、メタデータだけが検索基盤にも移る、連携するというものです。分かりづらくて申し訳ありません。

●矢吹 ありがとうございます。もう一つ質問があります。「国際ジャーナルなどに掲載された論文などの出版物と、それらに関連する研究データはどのよう

にリンクされるのでしょうか。ご存じでしたらご教示いただきたく存じます」。

●大波 現在、国際ジャーナルに載ったデータを流通させる基盤として、海外のデータ基盤であるCrossRefやDataCiteといった海外データ流通リポジトリと連携を行い、情報をCiNii Researchなどに取り込もうとしています。その海外由来のデータを使って国際ジャーナルのものや英語のものをつなげていき、そちらで成果物を確認できるようにしています。若干途上の部分もありますが、そういったポリシーで進めています。

●矢吹 ありがとうございます。さらに追加の質問です。「JAIRO Cloud以外のリポジトリとGakuNin RDMとの連携予定について伺えますでしょうか」という質問です。

●大波 JAIRO Cloud以外のリポジトリに、分野別リポジトリや、さまざまな大学固有のリポジトリがあるパターンもあります。したがって、画面では国立情報学研究所の基盤を使うことをお勧めしましたが、必ずしも使う必要はなく、別途これを迂回するようなルートを経てCiNii Researchで検索できるようにする方法でも大丈夫です。都度対応ということにはなりますが、大学や研究所とのご相談の上でこの流れに載せるということは十分可能ですので、ご理解いただければと思います。

●矢吹 ありがとうございます。もう一つ質問があります。「ご紹介されたメタデータのスキーマは、『2023年度までに、DMPと連動したメタデータの付与の仕組み』で述べられているメタデータの標準スキーマとになっていく予定なのでしょうか」というご質問です。

●大波 メタデータと連動していくかというところは内部でかなり議論があり、「そうだ」と言いづらいたところがあります。ただ、そうしていくことが、諸外国

NII RCOS

まとめ

- 『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』(2021年)に沿う、**研究データ管理が必要な状況**
- メタデータ共通項目にマッピングされた研究データは、**研究データ基盤システムNII RDC**(NII Research Data Cloud)
 - GakuNin RDM(管理基盤)
 - JAIRO Cloud(公開基盤)
 - CiNii Research(検索基盤)
 を利用して適切に管理し、**新たな知の循環へ**
- 2022年4月にCiNii ArticlesはCiNii Researchに統合

(図 11)

でも進められている機械可読の DMP (machine-actionable DMP) の利用の仕方ともつながっていき、単純にデータを登録するだけでなく、それがまた社会を良くするアウトプットが出せるものになっていくと考えて取り組んでいるので、大変良いご質問だと思うのですけれども、引き続き検討中という答えにさせていただきます。

●**矢吹** ありがとうございます。最後にもう一つ、「GakuNin RDM 上で DOI を取得した場合、リポジトリにデータを連携した場合にその DOI を引き継ぐのでしょうか。あるいは、GakuNin RDM とリポジトリのそれぞれで DOI を取得することになるのでしょうか」という質問です。

●**大波** まず、DOI は GakuNin RDM で取得できるのではなく、公開基盤の JAIRO Cloud で登録できるようになっています。一方で、DOI はそもそもジャパンリンクセンター (JaLC) や他の分野別データベース、あるいは海外のリポジトリで並行して取られてしまう場合もあり、複数の DOI を持った場合にどちらを優先させるかということは DOI の中で大きな議論になっていると思います。一つの答えとしては、最終的に CiNii Research で検索するときに名寄せをして複数の DOI を示すようにしています。そのような形で、DOI を複数取ってしまった場合でも、それを織り込む形で基盤の方で対応できるようにしています。

SPARC Japan セミナー2021

「研究データポリシーが目指すものとは」

総合討論(第1部)

研究データポリシーをめぐる論点整理を中心に

林 和弘 (科学技術・学術政策研究所)

池内 有為 (文教大学 文学部)

三宅 隆悟 (文部科学省 学術基盤整備室)

田野 俊一 (電気通信大学長)

大波 純一 (国立情報学研究所 オープンサイエンス基盤研究センター)



会場からの質問に対するパネリストの回答

●林 大変ありがたいことに質問をたくさん頂いてますので、前半はまずそれにお答えして、後半は今後の研究データポリシー関連活動にどうつなげていくかという議論をさせていただきたいと思います。

●池内 質問を読み上げさせていただきます。まずは純粋な疑問ということで、三宅様への質問です。「基本的な考え方において目指す主要な数値目標にある、2023年度までにデータマネジメントプラン(DMP)およびこれと連動したメタデータの付与を行う仕組みとは、具体的にどのようなもののでしょうか。2023年度から科学研究費補助金(科研費)と公的資金による研究においてDMPと研究データのメタデータ提出が必

須になるということでしょうか」という、現場としては大変気になる質問が来ています。

●三宅 こちらの内容については、各競争的資金制度などによって書きぶりは変わってくると思うので、その制度に合わせた求めをしていくことを考えています。基本的には、データマネジメントプランを作成して、それを基に管理することになるかと思っています。これについては、管理すべき研究でも変わってくるので、それぞれの競争的資金の中で詳細を定めていくことになるかと思います。もちろん、これもだんだん慣れてくるので、最初のうちはそのような内容から、少しずつ具体的な内容が入ってくると思っています。

科研費については、ただ今、検討していただい

ます。特に科研費は金額が大きく、件数も多くて、研究の規模も大小さまざまなため、かなり慎重に検討していると聞いています。具体的な中身につきましては、また新規公募分において内容を明らかにしていただく形になるので、それまでお待ちいただければと思います。

●池内 メタデータとは、機関リポジトリなどでの公開を前提としたときに付けるメタデータという理解でよろしいでしょうか。

●三宅 はい。説明が漏れていました。これもムーンショットの関係では先行して動いていて、イメージ的にこういう項目をとということ動いていたりもします。これもいろいろカスタマイズが入ると思いますが、研究データの公開・管理をする際に研究データに付与するデータ項目を明示して、それを付けた上で公開するということが研究費ごとに求められるというものです。

●池内 大波さんのご講演でも、メタデータについては検討が慎重に進められていると伺いましたが、現場の研究者や研究を支援する方たちにとっては、必要なものを何度も重複することなく、無駄なくきちんと入力できるようになると大変助かります。どうぞよろしくお願いたします。

続きまして、取り組みを実施していく上での疑問ということで、お三方にそれぞれ1点ずつ質問が来ています。

まず、引き続き三宅様への質問です。「オープンサイエンスや研究DXの取り組みは文部科学省だけではなく全省庁が一丸となって取り組む必要があると思いますが、省庁間の連携はどのようになっておりますでしょうか」。

●三宅 まさにご指摘のとおりで、今回ご紹介した科学技術基本計画や基本的な考え方は、実は内閣府主導でまとめられています。基本的な考え方も統合イノベ

ーション戦略会議で関係の大臣が出席している会議の中で決定され、政府としてはこれを全省的な取り組みとして推進していく形になっています。どうしても文部科学省の関係が非常に大きいため、われわれの関与が大きくなりますが、取り組み自体は全省を挙げてのものです。今回は研究データを中心に論じていますが、当然、民間のデータなども含めて連携していくという大方針もございます。また、もっと広げれば海外のデータ基盤との連携も含めて、さまざまな形で一丸となって取り組んでいくという方針で進めていますので、文部科学省だけでなく、各省を含めてさまざまな形で推進していきたいと考えています。

●池内 ありがとうございます。力強いお言葉だったと思います。

●林 田野学長、今の説明を受けて、大学の現場としては、文部科学省だけを向いているということもないと思いますし、先生の取り組みは厚生労働省や経済産業省などのさまざまなプロジェクトとかかっていると思います。先生側の立場から見て、省庁間連携について何かコメントを頂くことは可能でしょうか。

●田野 おっしゃるとおり、私の取り組みは文部科学省だけでなく、いろいろなところと関係しています。実際、内閣府に話しに行ったり、経済産業省の研究所に行きますが、縦割りですので、複数が関わっていてもなかなか「一緒にやろうね」とはならないのが問題です。「大学の問題なら文部科学省ではないか？」という話になるのですが、規制を改革してもらうために内閣府などに話しに行っても、なかなか腰が重かったりします。ですから、もっと仲良くなってほしいと思います。

●池内 今のお話からすると、データの話をつきかき、割とこういった話を通じやすくなっていく土壌ができてくるのかなという期待が少し生じました。あり

がございました。

引き続き田野学長への質問です。「東京 BPSD プロジェクトの取り組みを大変興味深く拝聴しました。実際にデータを集めるに当たって、センサーの設置などではプライバシーに関わるさまざまな制限や説明も必要とされるのではないかと思います。どのようにクリアされているのでしょうか」。

●**田野** それは大変苦労しています。まず、実験をするためには倫理委員会を通さなければなりません。この実験は（連携先の）医学部、介護系、データ管理部門の3カ所の倫理委員会を通してゴーサインが出ます。その際、相当な量の説明を書かされます。BPSD（認知症の行動・心理症状）であれば、患者一人一人に1項目ずつ説明したということをチェックしたものを付けて、ようやくデータがもらえます。BPSDに近い方々ですので、なかなか理解に時間がかかる方もいます。基本的には患者とその家族の方々に1人1時間ぐらいかけて一個一個説明しています。そのぐらいしないとデータがもらえません。さらに、データの利用目的も、BPSDの予測・対処に使うということに限定されており、それ以外での使用は禁止です。非常に大変です。

もう1年半ぐらい実験を行っていますが、やはり音声重要だということが分かってきました。音声といっても、ため息もそうですし、本当の言葉も重要です。それで1個センサーを増やそうと思うと、また倫理審査にかけて、もう1回説明が必要となります。このような手続きを経てようやく実験ができます。

ただ、患者や家族の方々はとても困っていますので、非常に協力的です。それ以外の方に、データを自由に使わせてと頼んだ場合、恐らくノーと言うのではないかと思います。BPSDは難しい問題だからこそ協力者は多いと思っています。

●**池内** 後続の機関なり研究者にとっても非常に示唆に富む事例だと思います。3カ所に倫理の申請をする

ということですが、それぞれ違った部分を審査されるのでしょうか。

●**田野** 同じ研究目標ですけれども、電気通信大学の倫理委員会は、データの管理や、どれぐらい保存しているのか、どれぐらい個人名を出さないかなどを審査します。私たちと連携している順天堂大学は、医学的に見てそのようなデータを取っていいのかを審査しますし、介護センターのようなところでは、介護士としてOKなのかという観点で審査します。立場の違いで見る項目が結構違っており、三つを合わせて全体の研究の倫理審査会としています。

●**池内** 非常に丁寧な手続きを取られていることがよく分かりました。ありがとうございました。

3点目は大波様への質問です。「GakuNin RDMでの研究データ管理ですが、企業などとの共同研究の際、GakuNin RDMを利用していない企業との研究データをどう管理するかなど、まだ使い勝手が良くないと思います。そのあたりは何かお考えになられているのでしょうか」。私の所属している大学も実はまだGakuNin RDMに入っていないので、個人的にも大変関心のあるところです。よろしく願いいたします。

●**大波** GakuNin RDMに登録していない方など、さまざまな方が共同研究にはいて、そういう人とGakuNin RDMを使う場合はどうしたらよいかというご質問だと思います。一つの答えとして、GakuNin RDMの登録情報については、外部のさまざまな公共リポジトリとの連携を行っているので、データそのものについては、そういったところとGakuNin RDM登録のアドオンを介してつなげるということをお試しただけだと思います。

その一方で、GakuNin RDMという基盤をプロジェクト管理として使おうとしたときに、そもそもログインもできないし、どういうものかも分からないというところで、さまざまな人が集まるプロジェクトで一緒

に取り組んでいくことが非常に難しいという話になると思います。それについては、GakuNin RDM の認証ルーティンをもう少し広いものにしていくべきではないか、フリーで使える IdP (Identity Provider) と呼ばれる認証機関を介して設定していくべきではないかといった議論も中でしているところです。まだどのようなことになるということは言えず、ご迷惑をお掛けしている部分もあるのですが、ご意見は受け取らせてさせていただきます。ありがとうございました。

●池内 もちろんセキュリティ第一ということは重々承知していますので、ご無理のない範囲で、しかしご対応を考えていただけるということで、ありがとうございます。

今度は評価の話について、三宅様への質問です。「研究開発を行う機関の責務として、研究データマネジメント人材支援体制の整備および評価とありますが、機関が自己評価を行うということでしょうか。大学の場合、具体的な評価の方法の想定はございますでしょうか」。

●三宅 詳しくは基本的な考え方の本文をご参照いただければと思いますが、大きく二つの観点で書かれています。

一つは、研究データマネジメントを行うことに対する支援が必要となりますので、支援体制の整備をまず機関の方に求めています。そのためには当然、そういう人材のスキルや取り組み自体が大学の中で職務として位置付けられなければいけません。それがどういう形かということについてはいろいろあり、また関連部署も非常に多くて、どこに置くのかということもあるかと思いますが、取り組み自体を職務として明確に位置付けて評価するというのがまず一つの観点だと思っています。

もう一つは、研究データを管理すること自体の研究者への評価という観点です。これについては基本計画の中でも、具体的な取り組みの状況について、研究者、

プログラム単位、機関など、いろいろな段階で評価体系に導入するという方針が示されています。この具体化については別途評価部局の方で検討しており、その内容を踏まえて、具体的にその機関の中でどのような形で行うのか、いろいろな段階があるけれども、その中でどうやって取り組むのかということを検討するという流れだと考えています。

●池内 今回は人材の話はメインテーマではないのですが、どうしても今後議論が必要になってくる部分かと思っています。ありがとうございました。頂いている質問は以上です。

●林 ありがとうございます。まだ質問は受け付けていますので、適宜 Slido の方にお願ひできればと思います。

討論

●林 本日のタイトルは「研究データポリシーが目指すものとは」です。多くの大学からご参加いただいています。研究データポリシーをどうつくればいいのか、あるいは、つくっただけではまずいよねということを理解されている方が、きっとこのセミナーに興味を持って参加されているのだと思います。その流れで、行政、大学の現場、データ基盤整備の三者三様の立場からのお話を頂いて、皆さんは「どうつながるの？ この話」とお感じになったところがあるのではないかと思います。あるいは、それぞれの間に認識や実践のギャップを感じた方、あるいは、それら三者と図書館なら図書館の方々とのギャップをすごく感じた方もいるのではないかと思います。全てにお答えできるとは思いませんが、この辺を、まずは研究データポリシーをめぐる論点ということで整理していきたいと思っています。

何より皆さんは、電気通信大学が研究データポリシーに関してどのように取り組んでいるのかということが一番聞きたいのではないかと思います。田野学長か

ら、電気通信大学のビジョン、あるいは大学や科学研究そのものがこう変わっていくことを示していただいた上で、大学の経営として研究データポリシーにどう取り組んでいるか、場合によっては2部の上野さんにも参加していただいでご説明いただければと思います。

●**田野** これは結構大変な質問です。本学のデータポリシーについては、図書館を中心とした体制で粛々と取り組んでいるところです。それはそれとして、私が考えているのは別のことです。研究データポリシーは使えますけれども、それだけで全部が進むとはあまり思っておりません。そのため、乖離があるのだらうと思います。

先ほど省庁で差があると話しましたが、それが典型的で、文部科学省の立場では、研究データをオープンにしないと論文は訴えられますし、エビデンスがないと言われますので、論文のデータをきちんと管理しないとまずいというスタンスです。それは管理であって、使うことをあまり考えていないと思います。

一方で、恐らく経済産業省や内閣府は、新しいものが組み合わせられて新しいイノベーションが起こらないようなデータは無駄だから、新しいイノベーションが起こるような仕組みをつくりなさいというスタンスだと思います。

文部科学省と経済産業省や内閣府は研究データに関する立場が違ってきます。大学等で作成された研究データポリシーを見ると、やはりきちんと正しくエビデンスとして蓄えておかなければならないというように見えます。それは重要なのですが、それでイノベーションが起こるかどうかは分かりません。先ほど、(三宅室長が紹介された NISTEP の実態調査の「研究データ公開のポジティブな影響」について) 良かった・悪かったという数値がありましたが、それでイノベーションが起こったというデータはありませんでした。本当は調査項目に新しいルールを発見した、規則を発見したというデータがあればいいのですが、そうではな

く、自分の研究に役立ったというデータしかないのです。ですから、イノベーションが起こることと、管理するということは少し違うように思います。それは両方一緒にやるというより、別々にやって統合するのがいいのかなと思っています。答えになっておりませんが。

●**林** いえいえ、正直、そこをすらすらと答えられる大学はまだ一つもないのではないかと思います。三宅さん、少しコメントしたいことがあるのではないのでしょうか。

●**三宅** 文部科学省においても、研究 DX に向けて研究データを管理して活用していくという方針で進めています。研究データ管理の話で先行して出たのが研究倫理や研究公正の話、10 年保存の話などです。そのような経緯もあり、そういうイメージから入られているのだと思いますが、文部科学省でも、他のデータリポジトリの話や各分野間のデータベースの話をつないで一体的に活用していくという方針は出していますので、そこは経済産業省なども含めて同じ方向を向いて取り組んでいるつもりです。それが足りないということであれば、さらに頑張りたいと思います。

その上で、やはり足りていないのが具体的なケースです。やったらどう役立つのかという話がなかなか見えてきませんので、取り組んだ方がいいに決まっているのだろうけれども、どうしようかなと思うのは当然のことだと思います。したがって、われわれとしては、マテリアルサイエンスを中心としたユースケースの取り組みに注力しているところです。

こういう取り組みは、例えば気象の分野や天体の分野など、そもそもそれをしなければ研究にならないという領域においては、DX の前の段階からいろいろとされています。しかし、恐らくこの話は、領域を超えてさまざまな研究領域においても同じような動きができるし、領域を超えたデータの活用も新しく見えてくると思います。確かにそこに関してはなかなかケース

が見えづらいところがあるので、そういうところをしっかり見せながら進めていかないと、なかなかこの取り組みは具体化していかないのではないかと考えています。そういうところに注力して進めていきたいと考えています。

●林 もしよろしければ、上野さんのお立場でどのような取り組みをしているかという情報も共有していただけでしょうか。

●上野 田野学長からも話がありましたが、私もこの取り組みには二つの方向性があるのだと認識しています。

一つは、先生方が論文発表等で研究データを管理しなければいけない一方で、そのオープンになっているデータを使っているいろいろなことをしたい。特に本学の図書館ですと、5年前からセンシングができる仕組みを設けて、人感センサーやCO₂センサー、カメラのデータ、個人情報に関わるデータなども取得できる設備を別途構築してきました。これを用いて、図書館のパフォーマンスの数値化や向上、あるいは教育分析などにつなげていきたいと考えています。その中では、やはり今あるデータを図書館だけで持っている意味がないので、学内の先生方や、できれば学外の方にも使ってもらえるような仕組みにして、先生方だけではなく図書館としても乗っかっていくのがいいのではないかと考えています。

それを踏まえて、文部科学省からの通知を受けて、昨年、図書館のある学術情報課、研究推進課、情報基盤センターに声を掛けて、ポリシーとシステムの準備に協力してもらって学内体制を組みました。これを研究担当理事に説明して、研究担当理事から役員にこの案件を進めますという話をしてプロジェクトが開始しました。現在、情報基盤センターでは、GakuNin RDMの利用申請を行い、準備中です。

一方で、ポリシーをどうするかという中で、私もまだ先生方とのコミュニケーションや勉強が不足してい

るところもあるので、先生方が実際にどのように研究データを管理しているかというアンケートを実施する準備をしているところです。これを年度明けぐらいまでに集計して、その結果を踏まえて、本学は管理プラス研究の推進という2点を実現するためにどのようなポリシーで進むのか、また、それを踏まえてどんなシステム構成にしていくのか、どんなサービスをするのかということを検討していきます。

●林 大変重要なお話で、この企画をやってよかったと思うのですが、最初に、通達があったからいきなりデータポリシーの検討を始めたのではなく、まず電気通信大学として研究の姿はこうあるべきということをきちんと認識した上で、ポリシーをどう役立てるか、そのために必要なステークホルダーは誰かという順序でデザインして進めているということで、このストーリーは多くの大学の参考になるのではないかと思います。

三宅さんの代わりに言っていますが、行政側も決してポリシーをつくるためにつくれと言っているわけではなく、行政なりに研究DXの姿を描き、それはアカデミアと全くイコールではないかもしれないけれども、それに向けてマネジメントに役立つポリシーづくりをしましょうという流れになっています。ですから、その辺にある公開されているものを少し直せばできてしまうというような対外的なポリシーは政策としては全く本望ではありません。その点で、何のためにつくるかということを中心に考えていくという姿を示していただいたのは大変重要だと思いました。

三宅さん、すみません、私が話してしまいましたが、何か補足コメントはありますか。

●三宅 まさにそのとおりで、何のためにしているのかということが一番大事です。例えば各研究機関や大学も、研究方針などの戦略に基づいて、どうしていくのかという階層構造になっていると思います。単純にここの部署がやればよいという世界ではなく、研究推

進や情報基盤、図書館などいろいろなところをまたがる話ですので、非常に難しいことはよく分かっているのですが、検討の中でそういうものを定めていただけると大変ありがたいです。

●林 その流れで、田野学長にもう一つ伺いたいことがあります。この手の話を進めるときは必ず抵抗勢力が生まれます。誰がいい、悪いではなくて、構造的な革新にはイナーシャ（慣性力）が働く、あるいは反対の動きが出るものだと思います。総合大学の場合、特に人文社会系を含む場合は、理工系に比べると閉じた世界のエコシステムで回っていればよいことも多く、成果を外に出す必然性が感じられないというコミュニティがあったりすると、その方々に納得してもらうのが非常に難しくなるのですが、電気通信大学の中にそういう抵抗勢力的なものは、そもそも構造的に存在し得るのでしょうか。

●田野 今のままのエビデンスを残すようなデータポリシーでは、公開しなさいと言われること自体、みんな嫌がるのです。それは人文社会系を含め、分野に関わらず、あまり使いたくないデータ、使われないデータをエビデンスにするために抛出しなさいというように見えてしまうと全教員が嫌がるでしょう。どう利用するかによって違うのです。利用されなさそうだけでもやらなければならないというように、上から来てそれを整備しようとする、公開するのが面倒で出したいくないということになるわけです。

では、データが非常に使えるものになったとします。データがみんなで使えるようになるという欧州ではきちんと公開しようとしています、そうすると次はどうなるかという、公開した自分が負けるかもしれないのです。トップ機関の素晴らしい研究データを、あまり分析されていないけれども公開すると、誰かがそのデータと別のデータを使って新しいものを生み出してノーベル賞を取るということもあり得るのです。

したがって、データを公開するならば、それがその後どう使われたかというトラッキングが重要です。それをきちんと保証しないと駄目だということにも欧州は対応しているはずですし、NIIでも研究しているのではないのでしょうか。これからは、論文を書いた人は表彰されるけれども、良いデータを公開した人も表彰されるべきだ、だから評価制度も変えないとならないということです。本当に使えるデータでしたら、それを一生懸命公開した人を表彰すべきですし、トラッキングシステムもつくるべきです。形式的にやっていると、面倒くさいといってみんな反対しますが、本当に使えるデータが公開され、活用されなければ意味がないので、公開した人のデメリットにならないようにきちんと保証されるようなシステムをつくる。公開するデータも、どこには出していいけれどもどこには出してはいけないというセキュリティをかける。そのようなことが次のステップとして重要だと思います。

●林 この辺のエンジニアリングは大波さんのところの真骨頂です。先ほどセキュリティ絡みで似たような議論がありましたが、田野学長がおっしゃった課題を解決するような兆し、使っている人たちはこんな活動をし始めたといった話、あるいはエンジニアリング的にはこういう考えがあるというコメントを頂けますか。

●大波 例えばデータを評価したり登録・公開したりすることがどのように社会を良い方向に動かしていくかということについては、いろいろなサービスを公開することで、こういった使い方もできるし、ああいった使い方もできるだろうという提案はいろいろあります。例えば資金配分機関で絞り、大学ごとにどんなプロジェクトが取れるかという統計情報のようなものをどんどん出していき、どこの部署が日本の中ですごく頑張っているかということを見えるための基盤として、まずは情報を見てくださいということもあり得る話だと思います。

その一方で、分野によって、そういうアウトプット

が出しやすい分野と出しにくい分野があり、細かい論文をどんどん出していけるところと、論文や本を一生に1度ぐらいしか出せないところを一律に見てしまってよいのかということもあります。それはまた基盤と行政側でどのような形が理想かということを話し合っ決めて決めなければいけないと思います。そういうところに役立つテクニックを出していこうとわれわれの方では考えています。

●林 その意味では、管理基盤の上のサービスレイヤーの設計がすごく面白いというか、田野学長が示された課題や評価の課題も、理想的には全部その中間プラットフォームで解決できるはずと思いながらみんな取り組んでいるとは思いますが、どこから始めるかという鶏と卵問題もありますが、隗より始めよということで、良く言えばカンブリア紀のようにいろいろな取り組みが無数に出てきて、残ったものが勝つという見方もできる現状であり、そうなるとう数が多くの方がいいということはあるかもしれません。大学にはそれぞれ性格があるので、データ管理基盤の上のサービスレイヤーは十人十色になると思います。この部分が研究戦略になり、その研究戦略に基づいたデータポリシー作成になっていくのだらうと思います。

なかなか悩ましいところも多いですが、前半のこの討論では、実務というよりは、そもそも論的なところもう少し深い議論ができるかと思っています。逆に、パネルの皆さまから投げ掛けたい論点、疑問、あるいは要望などがございましたらぜひお願いします。

●田野 おっしゃるように、データをどのように使うかということでデータポリシーをつくらなければならないと思います。今のデータポリシーはかなり管理に重点を置いているというか、公開しないと世の中から叱られてしまうというものですが、本当は応用しようという感覚でデータポリシーを作成すべきだと思います。そして、それは国がやらないといけません。トップレベルの実験をしている人たちは公開したくはな

いです。しかし文部科学省は、「公的研究費が配分されているのだから説明責任があるだろう。データを公開しなさい」と言います。しかし、公開したことがきちんと評価されないのであれば、実験データの全てではなく、ある特定のデータだけを公開したりするわけです。そうすると本末転倒です。

データをみんなで共有して新しいことに使おうと言っているのですが、一方で、きちんと研究しているというエビデンスのために公開しようとするという二つの流れがあります。ある国が「使う」という制度を設定すると、その国は非常に進んでいきます。その国の人たちは、全部のデータを公開して、みんなで共有して新しい発見をするようになります。しかし、別の国では全然違うポリシーがあり、「きちんと管理しましょう、どう使うかはあなたの勝手です」というポリシーだとすれば、その国は廃れていきます。

ですから、ポリシーというのは国が栄えるか栄えないかに関わる大変大きな問題だと思っています。欧州は結構真面目にデータをオープンにして、重要なのは論文数ではない、データを公開するあなたは偉いというように、データ公開に関する評価を全く変える動きも出ているそうです。日本もそれぐらいやらないといけません。今のままでは管理のためのポリシーに見えてしまいます。トップダウンで検討した方がいいと思っています。

●三宅 どのように活用していくかとともに、どのように公開していくかということも考えないといけないと思っています。もちろん公開が善というわけではなく、共有すべきところに共有するという管理の仕方も重要です。全て公開しなさいと言われてればディフェンシブになるのは当たり前なので、理解してメリットを感じていただいた上で共有していく。それはオープンにすると意味ではなく、必要な相手と必要な情報交換ができるというところを押さえて仕組みをつくっていく必要があると思っています。

最初の議論がオープンアクセスという言葉から始ま

ったので、全てを公開すると言っているように聞こえて、そこで少し反発があったという経緯はあるかと思えます。そうではなくて、それを行うことによって研究が進むのだという理解を進めていくことが重要だと思っています。

研究領域によって事情が異なるため、なかなか一律に全て同じルールでというのが難しいことは分かっています。そこについてはわれわれも各分野と連携して、例えば民間との連携が非常に強い領域では、特に民間側はそれがまさに将来の種になることから、公開する、公開しないということには非常にシビアな議論があります。メリットを一つ一つクリアにしながら仕組みを整えていくことが必要だと思えます。なかなか一朝一夕ではできませんが、国としても、一つずつ理解を頂きながら進める必要があります。国全体、産業界、学全体も含めて、機運を醸成しながら進めていくことが必要だと思っています。

●林 政策的には、昔、産学連携を進めましょうと言って、大学と企業の共同研究あるいは特許の数を数値目標にしたら、田野学長が懸念されたとおりの形式的な数だけ増えて、見かけ上、産学連携は進んだのだけれども、1件当たりの額が非常に小さく、特許は取るけれども、それが産業構造を変えたり産業に飛躍的な利益をもたらしたりしたかということ、全然そんなことはなかった時期があったという苦い歴史が存在します。研究データの在り方がそれと同じことになったら、本当に日本の国力そのものが下がりがねないイシューだと思います。非常に重要なご指摘を頂いたように思います。

逆に三宅さんから見て、現場の声という意味で何か言っておきたいことがあれば、個人的なことでも構いませんのでお願いします。

●三宅 やはりこういう検討をしていく上で、具体化していく中での隘路というのが一番大事な問題だと思っています。田野学長からも、こういうところで苦労

したというお話を聞いて非常に面白かったのですが、大学が個々で解決する課題と、国全体で解決した方が早い課題、それぞれの役割に応じて対処していかなければいけないと思います。例えば、こういうところに関して方針が示されているとやりやすいという観点で、もしご示唆いただけるものがあれば教えていただければと思います。

●田野 政府にお願いしたいことはいろいろあります。「管理する」と「使う」のうち、「使う」の方向ですと、いろいろな研究をしなければなりませんから、それを政策も含めて国家プロジェクトとして進めるべきだと思います。政策というのは、公的研究費を使った研究は全部データを公開するというのもそうですし、それ以外にも、いろいろなところでデータが集まってきているので、それも公開しないと駄目ですよというものを本当はつくりたいのです。

卑近な例で言うと、例えば部屋にエアコンがあって、いろいろなところにセンサーが付いています。最近では、CO₂が上がると換気が悪いということで、CO₂センサー対応のエアコンが出てきていますが、これを導入しようとするとな数千万円かかります。しかし、CO₂センサー自体は1,000円程度です。データをお互いにリンクできれば全然問題ないのです。例えば、国が「社会システムの中でいろいろなセンサーがあるでしょう。機能もいろいろあるでしょう。そういうセンサーも機能も全て情報をオープンにしないといけない。セキュリティはきちんと政府がつくったシステムで守っています」というようにすると、データがオープンになり、機能がオープンになり、アクセスを許された人たちが連携を取って、古いエアコンも1,000円で簡単にCO₂対応のエアコンにできるわけです。そういう政策をたくさんつくってほしいと思っています。

例えば霞が関に入れるシステムはそうしなさい、国立大学に入れるシステムはそうしなさいというように、小さいところから始められます。そうすれば、「公的研究費を使った研究はオープンにしなさい」と言うよ

りも、徐々に広げていけるのではないかと思います。

お金がゼロ円で済むように政策をうまく使って、みんながデータをどんどん使えるような世界を目指してほしいと思っています。使うという立場で考えると、いろいろな政策が考えられるのではないかと思います。

●林 一貫して使うという立場から、ある意味、バックキャスト的に使う側から考えていくという本日のメッセージが大変強く伝わってきました。使われてなんぼのデータだという話も含めて、全くそのとおりだと思います。使われてなんぼのときに、では日本国としてどのような研究データの使われ方をしたいのかというビジョン、大げさに言うと、日本国としての研究戦略性をきちんと示しましょうという話にもなってくるのだと理解しています。

●田野 そのときに、さすが日本だという政策にしてほしいのです。厳しく中央管理をしてがちがちになっているようなデータ管理ではなく、さすが日本だという政策にしてもらえるといいなと思います。そうすると、いろいろな人が日本に集まってきて分析を始めるのです。これも競争で、全世界がそうなってしまうと日本は負けていきます。ですから、さすが日本だなどというものがほしいと思っています。

●林 本日の論点ではないのですが、まさに国際性のところだけでも半日ぐらいシンポジウムができるかと思えます。留学生や研究者に来てもらうことを含めて、アカデミアとしても産業としても魅力的な市場であるためには、研究データインフラが国際的に見て魅力的でないと話にならない。まず話はそこからだという時代が来るのが分かっています。では、そのために何をいつ、どこまで、どうやってするのか。日本の戦略として、少なくとも損はしない、できれば日本のプレゼンスが上がる、ついでに言えばアジアの中のリーダーシップを取れる形でどのように進めていくかという議論なのだと思えます。

●池内 これに関しては、「データポリシーは管理のためのものではない」という田野学長のご意見に賛同します。研究データが公開・整備されている国や研究機関は、世界の研究者から尊敬を集めると思います。日本が国としてそういった余裕を持った考え方ができるようになることを望みます」という意見が会場の方からも届いています。

●林 ありがとうございます。

●池内 三宅室長が紹介された研究データ公開のポジティブな影響ということでも、やはり実際に公開している研究者の方々から、研究データを公開して国際的な共同研究につながった、研究への注目が集まった、ビジビリティが上がったといった声もあります。まだ数は多くないかもしれませんが、そういった研究データ公開によって起きた良いことの共有は本当に重要だと考えます。

●林 これは無茶振りな質問ですが、大波さん、研究データポリシーを大学が策定するに当たり、どうサポートできるかということを考えたことはありますか。単にインフラをつくることも大事ですが、その上のソフトウェアやサービスなども切っても切り離せない問題で、その辺でうまい相乗効果を出せないのだからかということをお無邪気に思いますし、そうすると、NII も含めて加速度的に取り組みが広がる可能性もあるのではないかと思います。

●大波 NII のオープンサイエンス基盤研究センターの中で基盤づくりをしていて、海外の事例などの調査も日々行っています。現在、NII の中では、データポリシーをどうつくるかという検討を進めていて、自分の組織でつくるだけではなく、どうやったら他の組織に参考にしてもらえるような内容になるかということも、海外の状況も踏まえて盛り込んでいるところです。ですから、今から完全に予告できるものではありません。

んが、もしNIIの方でデータポリシーがつけられたら公開させていただき、参考にさせていただくのがまずは第一歩としてよいのではないかと考えています。

●林 そのときはNIIの喜連川所長にまず研究戦略のお話を頂かないといけないかもしれませんね。

●大波 そうですね。うまく皆さんと共有していければと思います。

●池内 本日のお話を伺っていて、田野学長がおっしゃるように、研究サイドからも政策サイドからも、結局目指す究極のゴールは、分野を超えて日本のデータが広く世界中からアクセスされて、より良く使われて、それによって科学が進展し、社会がもっと良くなることだと思います。そのためには、公開したデータを探すことができ、中身が理解できるようしっかりとメタデータが付いている必要があります。その部分でのNIIの役割は非常に大きくて期待するところですし、その先のデータが引用されたかどうかというトラッキングの部分についても、もし実装が進めば、評価にもつながるような非常に重要な指標を提供していただけるのではないかと考えています。

●大波 池内さんの言うとおりのロードマップを描いていきたいと思います。

●林 話しているうちに、あっという間にあと5分ほどになってしまいました。理念、概論、そもそも論的など、あるいはビジョンについて良い議論ができたと思います。

池内さんは研究データのモニタリングをずっとされていて、私もお手伝いさせていただいていますが、データを公開している数値を追いかけることにどれだけ意味があるのかという議論があって、田野学長の「どう変わったのか」というメッセージがやはり強いですね。

●池内 それが今回は本当に響きました。

●林 そこをどう見るか。有り体に言えば指標の開発になりますし、それが結局、研究インパクト計量からの研究評価にもつながるし、あるいは産業を興すためのKPI（重要業績評価指標）など、いろいろな話になります。時期尚早かもしれませんが、この辺も将来的には追いかけていかないとはいけませんね。

●池内 そうですね。現在の研究データ公開の進み方は、調査によると、やはり雑誌が主導していて、ジャーナルにデータを補足資料として付けなければいけないためにデータ公開が進んでいる状況です。そのレベルにおいては、日本の研究者も当然海外誌に投稿しますので、横並びとまでは言えないかもしれませんが、特に日本が遅れている感じではありません。ただ、いわゆるトップダウンでジャーナルが公開しなさいと言ったから公開しているというその先の、田野学長がおっしゃるような、研究をもっと回すためのデータ公開をどこまで日本が先陣を切ってやれるかというのが、これからポリシーを検討する上で考えていかなければいけない論点ではないかと思っています。

今回、電気通信大学のお話を率直にさせていただいて、田野学長の研究を中心とした部門と、データポリシーを考えているところが、少し離れてしまっているような感じがしました。ポリシーをつくる部門はつくる部門、研究者は研究者ということだと、研究者に「それならデータを公開した方がいい」と腹落ちしてもらえようなデータポリシーをつくることは難しいのではないかと感じました。率直に言うと、研究者の方がデータをどう使うかということをよく考えていると思いますので、研究者に「こんなことができるといいな」というところまで含めて話を伺える機会があると、ポリシーも実のあるものになるのではないかと感じました。

●林 ありがとうございます。大変重要なお指摘だと

思います。

●**田野** そういう感覚で、今、文部科学省がきちんと進めています。マテリアル実験のものをまとめる取り組みがスタートしているので、その人たちに、他の実験をどう使うかということまで考えて進めてくださいと指示すると、いい方向に進んでいくのではないかとと思っています。本学も、マテリアル系の10年ぐらいのスパンのプロジェクトの一つ入っていますが、他の人たちも使うようにするにはどうしたらいいかということのを常に考えてプロジェクトを回してくださるといいと思います。今はまだ「公開しましょう、公開しましょう」と言っているだけで、使うという観点でフォーマットやポリシーを決めて、それで指示して考えるのも一つの良いアプローチではないかと思っています。研究者が考えないとやはり駄目です。

●**林** ありがとうございます。重要な議論をたくさんすることができました。

私は研究開発法人のデータポリシーのガイドラインを策定に関わったときに、「研究機関のミッションに立ち戻る機会である。データの話ではなく、その機関が何の研究を何のためにするのかということ振り返る機会になる」ということをよくお話しさせていただいていました。大学ランキングの定規の上に無理矢理乗せられているきらいがある中で、むしろ自分たちの大学の強みは何なのかということを考えることで、それが研究戦略となり、データ戦略となって、ポリシーに落ちるのだということを私自身も本日は改めて確認しました。

3人のパネリストの方々からの情報提供とディスカッション、そして会場の皆さまのご質問に感謝申し上げます。それでは、これで第1部の総合討論を終了いたします。ご清聴ありがとうございました。

SPARC Japan セミナー2021

「研究データポリシーが目指すものとは」

総合討論(第2部)

研究データに関わる各ステークホルダーとの議論

- 林 和弘** (科学技術・学術政策研究所)
- 池内 有為** (文教大学 文学部)
- 三宅 隆悟** (文部科学省 学術基盤整備室)
- 田野 俊一** (電気通信大学長)
- 大波 純一** (国立情報学研究所 オープンサイエンス基盤研究センター)
- 能勢 正仁** (名古屋大学)
- 松原 茂樹** (名古屋大学 情報戦略室)
- 白井 知子** (国立環境研究所)
- 上野 友稔** (電気通信大学)
- 安原 通代** (国立情報学研究所)
- 八塚 茂** (バイオサイエンスデータベースセンター)
- 矢吹 命大** (横浜国立大学)
- 山形 知実** (北海道大学)



研究データポリシーをめぐる論点についてのコメント

●池内 討論に入る前に、第2部よりご登壇される名古屋大学の能勢様、名古屋大学の松原様、国立環境研究所の白井様より、研究データポリシーをめぐる論点について、簡単にコメントを頂戴したいと思います。

●能勢 私の本業は研究者で、宇宙空間で起こっている現象の研究をしています。実際にフィールドへ出て自分でデータを取って、そのデータを解析したりします。その他にGPS衛星などが飛翔している地球近傍領域について、人工衛星が観測したデータを使い、共同研究をしながら研究を進めています。ですから、まさに研究データのポリシーができたら、それに従って研究データの公開をする立場にいます。

研究者の同僚などの反応を見ていると、やはり研究データポリシーが策定されてデータを公開するとなると、「やらされている」「余分な仕事だ」と感じる傾向にあると思います。それは研究者の置かれている最近の状況が、非常に人が減っていることに呼応して業務量が増え、競争的研究資金の配分が大きくなって、大学に配分される研究費もその業績に応じて変わるという厳しいものになってきているので、さらに余分な仕事をするのかというのが現場の研究者の大多数の正直な感想になるのだと思います。

しかし、研究者がデータを公開しないと研究データの利活用は始まらないので、義務でやらされていると感じているというのは非常に残念なことです。研究デ

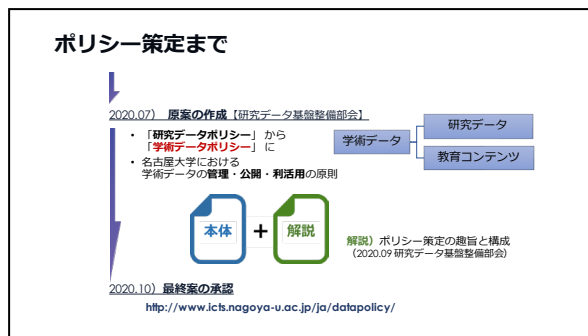
ータポリシーを策定する際に、どのように策定すると研究者がなるべく幸せにデータをシェアして公開していけるのかという立場で、名古屋大学の松原先生と一緒に、名古屋大学における研究データポリシーについて議論させていただいているところです。

●松原 私は情報戦略室に所属しており、大学の研究データ管理には情報側から携わっています。話題提供ということで一つの事例をお話しさせていただきます。

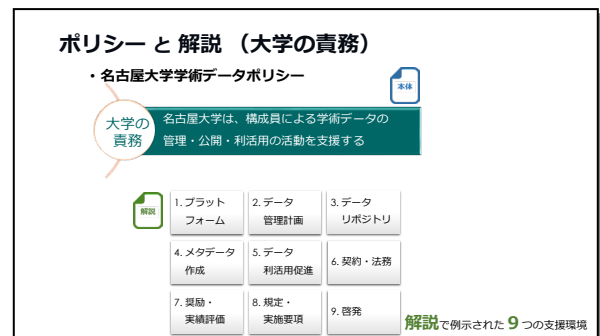
名古屋大学は2020年にデータポリシーの策定を行いました(図1)。2020年7月に原案を作成し、10月に最終案が承認されています。原案を作成するときにはいろいろ議論があったのですが、教育・研究を担う組織として、研究データだけではなく、教育コンテンツも管理・公開・利活用を積極的に進めるということを含めて学術データのポリシーとしてまとめました。

ポリシー自体は5項目から成っているのですが、それに解説を添えて、ポリシー策定の趣旨と構成を解説しています。このポリシーの5項目目に「大学の責務」として、構成員の活動を支援するということを定めています(図2)。それに対応して解説では、支援環境の例を九つ示しています。

それを受けて、データポリシーにおける大学の責務を遂行する主体として、実務を担当するマルチステークホルダーから成る組織をつくって活動を進めています(図3)。セミナータイトルの「研究データポリシーが目指すもの」という点では、このデータポリシーがまさに組織における研究データ管理・公開・利活用の力になっていると思いますので、データポリシーを



(図1)



(図2)

力に変えていくということが一つの論点としてあるのではないかと考えています。

●白井 私は能勢先生と同じように研究をバックグラウンドとしています。地球環境の部門にいて、研究の他に研究データの利活用、基盤の構築、運用などを行っています。

2018年9月のSPARC Japanセミナーで、国立環境研究所のデータポリシーをつくったときの経緯などを発表していますので、そのあたりについてはもう一度ご説明することはしませんが、国立環境研究所では2017年にデータポリシーを出しています。内閣府のガイドラインができる前だったのですが、これは研究所としてデータを積極的に公開していくという大枠の方針を示すデータポリシーになっています。

ただし、ポリシーは組織の理念や方針を示すものとしてつくる価値がある一方で、その中でどの程度実効性を持たせるかがポイントとなってくると思っています。国立環境研究所の場合、研究分野が多岐にわたっているため、具体的なデータ公開に関する取り決めなどはポリシーに盛り込まず、ある程度、現場で調節できるような形のポリシーにしてあります。それは現場としては大変ありがたくてやりやすい一方で、何らかの決まりによって決めたいというときに決めることができません。分野によって、自分で適切な判断ができる人材がいる、あるいは分野として非常に決まったマナーがあるという場合はいいのですが、そうでないと、全然決められない、判断できないということにもなってしまうという、いい面と悪い面があります。ですの

で、組織の中でどの程度細かいところまで決めるかということ、ある程度論点になるのかなと考えています。

あとは、ポリシーを決めるといっても、それはそれで専門的な知識が必要であり、組織としてどの程度支援できるかという問題がありますので、そのあたりも含めて、やはり一つ一つの機関が自機関の状況に合わせてつくっていくものだと感じています。

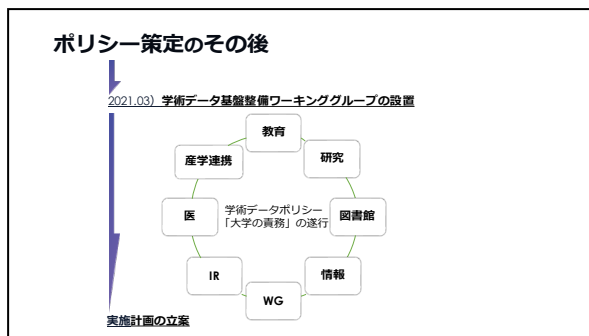
会場からの質問に対するパネリストの回答

●池内 それでは討論を進めたいと思いますが、ディスカッションに先立ちまして、まずはお寄せいただいた質問にお答えしていこうと思います。

●林 質問を紹介させていただきます。まず、回答講演者指定なしということで、「本学ではオープンアクセスポリシーは未策定ですが、2025年までに研究データポリシーの策定を迫られています。手順としては、まずオープンアクセスポリシー策定後に研究データポリシーの策定に着手すべきでしょうか」という質問です。

●池内 これはどなたか、いかがでしょうか。

●林 科学技術・学術政策研究所（NISTEP）で研究データポリシーに関する講演会をしたときに、その議論で大いに盛り上がりました。そのとき、講師としてNIIの船守さんに来ていただいたのですが、特に規模の小さい大学の場合は、まずオープンアクセスポリシーをつくる、ないしはオープンアクセスポリシーをつくっているところはそれを拡張して研究データポリシーにするところから始めるのが現実的ではないかという議論になりました。そのとき三宅さんもいらっしゃいましたね。私が答えてはいけないのですが、特にご意見がなかったのもまずは情報共有させていただきました。



(図 3)

もし回答がございましたら適宜受けたいと思います
が、ひとまず次の質問に移らせていただきます。

「研究データの提出・利活用に関して質問です。研究データポリシーに基づいてリポジトリ等で公開される研究データの権利処理はどのような形があるでしょうか。論文等の場合はクリエイティブ・コモンズ・ライセンスを付与しての公開も増えていますが、研究データの場合に適したライセンス処理の実例はございますか」。これは既にポリシーをつくって運用している研究機関からのご回答があると大変ありがたく思います。白井さん、どうですか。

●**白井** ライセンスの問題はよく質問を受けます。その都度、私も結構悩むのですが、地球環境のデータベースではオリジナルの利用規約をつくってもらうことが多いです。元々、いろいろ細かく決めたい人が多かったということもあり、それぞれのデータに合わせて自分で判断して決めてもらった方が後々いいので、ひな形の中で、これは違うと思ったところを変える形でオリジナルのライセンスを作ってもらうことが多いです。最近、割とデータジャーナル、あるいは普通の論文でも、ポリシーとしてデータを CC BY で公開してほしいということを言われることも多くなっています。そういったことにももちろん対応しているのですが、自分のデータはうるさいことを言わないからどんどん使ってくれという人は、CC BY などを出しています。

あとは、研究データ利活用協議会（RDUF）の中でライセンスの研究データライセンス小委員会が「研究データの公開・利用条件指定ガイドライン」をつくっています。この小委員会には池内さんも入っておられますが、やはりそういった文書があると、こちらとしても「こういったものもご活用ください」という感じで渡せるので、あれはつくっていただいて本当によかったと思っています。

●**池内** 研究データ利活用協議会でのライセンスのガイドラインの策定についてはいろいろと研究した経過

があって、クリエイティブ・コモンズ・ライセンスがやはり一番知名度も高くよく使われているということで、当初はそれを付けて出せばいいのではないかと議論をしていました。ただ、そもそもの話として、クリエイティブ・コモンズ・ライセンスは著作権があるものに対して付けるのですが、研究データは一般的に著作物として認められない場合が多く、しかもその判断が非常に難しいです。そのため、クリエイティブ・コモンズ・ライセンス相当というのでしょうか、名前としてはほぼ一緒で、例えば先ほどの CC BY のようにライセンスを表示する形で付けるのですが、クリエイティブ・コモンズ・ライセンスそのものではない方がいいというガイドラインになっています。後ほどガイドラインをチャットに貼りますので、よろしければご覧になってください（研究データの公開・利用条件指定ガイドライン https://doi.org/10.11502/rduf_license_guideline）。

●**林** それでは次の質問に移ります。「利活用のための研究データ管理、公開、共有について、教員からもメリットが理解できないという指摘を多く頂いています。説明しようにも、現場の事務レベルでは具体的な経験がないため、実感を伴う説明ができず苦慮しています。先行事例などの積極的な紹介をお願いしたいです」。このコメントは、これまで進められた方はどういう事例をお持ちですかという質問に読み替えさせていただきますが、いかがでしょうか。

●**能勢** 最近の例ですが、卒業生の研究を指導していて、卒業生や修士課程の学生の研究でもデータが出てくるわけですが、それを死蔵させるのは非常にもったいないので、名古屋大学の図書館に公開という形で置いたことがあります。そうすると、データを公開することによってやはり別の研究者に利用され、その方から「こういうデータを使ったのだけれども、一緒に共同研究しないか」という反応がありました。公開することによって利用されて将来的な共同研究につながる

ということは本当に実感するところです。

ただ、将来的には、やはりデータの公開に対して客観的に評価するような指標と仕組みが必要になってくると考えています。

●大波 成功事例等の積極的な紹介という点で、今でないとできないことは、やはり新型コロナウイルス感染症（COVID-19）対応の情報だと思います。パンデミックが始まって COVID-19 の情報がどんどん出てきて、それを基にワクチンを作ったり、COVID-19 にはどういう特徴があるかということを経済中の研究者がこぞって研究しています。そのスタート地点として、遺伝子のデータなど生物のデータは元々オープンにしてライセンスをはっきりして共有するという文化が十分育っていた面があって、そのおかげで COVID-19 のことが迅速に分かって、ある程度対応できているという状況だと思います。これがもし 20 年前など、あまり遺伝子の情報を公開するという機運が醸成されていない状況だったら、全然スピード感が違ったと思いますし、もっと世の中は混乱していたのではないかと思います。

そういう意味では、先にデータを共有してライセンスやポリシーをしっかりと定めておくということは重要です。生物のデータも、共有を始めた頃は皆さんやはり苦労があって、ここまでやらなくていいのではないかと思います。やっていたことが、ようやく今になって芽吹いたわけですから、先んじてやっておくことには意味があります。今回、日本全国で大学などに研究データポリシーを策定していただくに当たっては、将来何らかの研究の情報が急に必要になったときに、その状況になってからルールを定めて取り組むのでは遅いので、あるべき論として先に定めておくのがいいのではないかと説明の仕方ができるのではないのでしょうか。

●林 ありがとうございます。それでは、もう一つ質問を紹介したいと思います。「オープンアクセス方

針、オープンアクセス実施要領、研究データポリシー、機関リポジトリ運用指針と、たくさんの方針や指針を次々と策定することが求められ、猥雑になっていくように思いますが、研究成果の公開という一つのことをするに当たって、例えば OA 方針実施要領と機関リポジトリ運用指針を一つにまとめるなどした場合、何か不都合があるのでしょうか」という質問です。これはぜひ名古屋大学にお答えいただけるとういかと思いますが、いかがでしょうか。

●松原 多分そういった整理は必要になるのではないかと思います。例えば、本学では学術データポリシーと呼んでいますが、そういったデータポリシーの成立を受けて、機関リポジトリのポリシーが変わるなど、ポリシー間関係も変わっていくと思うので、統合などの整理の仕方は今後出てくるのではないかと思います。重要な意見だと思います。ありがとうございます。

●林 それでは、次の質問でいったん質問コーナーを終わりにしたいと思います。「大手海外出版社の学術誌ではデータシェアリングポリシーが挙げられ、その中では、Data Availability Statement として、データの要件や公開元の明示の仕方について提示が見られます。データポリシーはこれに対応できるものであればと思いますが、いかがでしょうか」。これも、既に策定された方々がどのように運用しているかという情報をシェアしていただくと助かります。

●能勢 研究データポリシーと学術出版誌のステートメントは基本的に同じようなものではないかと考えます。これらは、プリンシプル（原理）であって、研究者は、ではそれをどう実現するのかという点で非常に困っています。データを公開することには賛同するのですが、どのように公開するのか、メタデータにしてもどのような情報が必要なのか、出版社に受け付けてもらうにはどのようなデータをどこに置けばいいのかが

分からないということがあります。ですので、研究データポリシーや出版社のステートメントのもう少し下のレベル、実際にそれをどう実装するのかという段階で、具体的な手順の説明が必要だと思っています。

●林 雑誌というより、共同研究をするときの方が大変になりそうな気がします。特に国際共同研究になったとき、あるいは研究室同士で少しカルチャーが違うときに、データをどう管理するか、どう公開するか、どう共有するかというのは悩ましい話だと思いました。質問は以上とさせていただきます。

●池内 ありがとうございます。

討論

●池内 今の質疑の中からもいろいろな論点が出てきました。まず白井さんのお話の中で、2017年にポリシーを策定して、その後どう実効性を持たせるかというご発言があったと思います。データを公開しようというポリシーをつくった後に、その強制力がそれほど強いものでなく、しかしやはりデータ公開を進めていきたいというときに、どんな働き掛けができるかということが気になります。能勢先生のお話の中では、学生のデータを図書館のリポジトリに登録するという事例がありましたが、何か具体的な働き掛けや取り組みがあれば紹介していただけますか。

●白井 言いたかったのは、要するにデータポリシーをつくったからといって現場側がデータを出していくことにはつながらないということです。先ほどの質問にあったように、現場のモチベーションをどうやって上げるかは課題だと思います。研究者に対し「ご自分のデータがこういうふうに使われるかもしれないよ」と言って回るのはあまり得策ではない気がします。「あなたに言われなくても、私のデータのことは私が一番知っています」という反応をされてしまうでしょ

う。

人間というのは現金なもので、やはり自分の業績になるとか、お金をもらえるとかがあけると行動するものだと何となく実感しています。論文を出そうと思ったときに、データをそれなりのリポジトリから公開していないと査読にも回らない、あるいはそこまで厳しくなくても、パブリッシュするときまでには一般に公開できるようにしておいてねという学術誌・出版社は本当に増えてきています。やはり研究者は論文が出ないと業績にならないので必死で出します。論文がやっと書けたのに、データ公開で手間取ってサブミッションが遅れてしまったり査読開始が遅れてしまったりすると悔しいので、今度からは前もってデータを公開する準備をしておこうというようになります。そういういいサイクルが回っていく様子を実際に目の前で見てるので、やはり理念だけではなくて、実際のメリットが感じられるとうまくいくのだと思います。

もちろん、データが使われることによって共同研究が始まるなど、本当の意味でのいいこともあるのですが、それをぶら下げられても、天秤にかけるとやはり日々の忙しさに負けてしまいます。ですから、まずは目の前のメリットを感じて、進めていくうちに本当の意味でのいいことも時々入ってきて、さらにメリットを感じるというのがいいのかなと思っています。

●池内 率直なご意見をありがとうございます。研究者の実感として、本当にそうだと思います。バイオの分野でも、まずはデータを登録してからということに進んだ面が大きいと思います。

そうすると、ジャーナル以外の評価をしてもらえる可能性があるかどうか気がなるところです。大学、あるいは文部科学省でもいいと思いますが、データを公開することを、例えば論文を出版することと同じように評価しようというお考えはおありでしょうか。この件は、大学側はまず田野学長に伺いたいですが、例えば大学の教員の評価なり1年間の業績報告なりの中に、データが入ってくるという可能性はあるでしょうか。

●**田野** この取り組みがうまくいけば、それはもちろんあり得るというか、私たちが最初になければならないのだろーと思ひます。そのときに、単にデータを公開したというだけではなく、うまく使われたという指標が本当はほしいのです。それがないと、どんどん変なデータがアップされても仕方がないと思ひます。

●**池内** 「何でもいいから取りあえず」ではなくてということですね。

●**田野** 私たちは、とにかくデータを公開したら少しは「偉いよ」と評価しますけれども、それが論文1本と比較するとなると、研究者から見ても、あまりいい評価をされていないと感じるのではないのでしょうか。いろいろな研究者がデータを使ってくれたという指標と、1回しか使われなかったけれども非常に大きな発展に結び付いたという指標が本当はほしいです。この二つをきちんと評価しなければ、なかなかデータは公開されません。ただ、その評価の方法は一大学では開発できません。

5年ぐらい前に、カーキーというシステムがはまりました。これは、みんなでアイデアを出して新しい製品を作ることを進めるために、誰が発言して製品化までされたかをトラッキングするシステムです。本当に素晴らしい意見を言った人には例えば売り上げの5%が行くということで、みんな一生懸命話したり、データを公開したりします。そういうシステムがないと、大学が「幾つ公開したから1.5点ぐらいあげますよ」という話をしても、なかなか進まない気がします。

●**池内** そうですね。データの活用という話になると、もちろん研究に使われることもあれば、企業や市民に使われることもあり、いろいろな広がりがあるので、そのあたりのトラッキングの難しさもあるのかなと感じました。

三宅さんはいかがでしょう。

●**三宅** ご指摘の点はまさに考えなければいけない点です。科学技術基本計画上も、それについて取り組むべしということで、政府においても方針が示されています。具体的には、研究データの管理・利活用に関する取り組みをさらに促す観点から、これらの取り組み状況を研究者、プログラム、機関等の評価体系に導入するという方針が掲げられており、評価部局を中心に具体化に向けて検討が行われている状況です。研究データの管理・利活用の取り組みをどう評価していくのかということが議論されていて、その結果が国などの立場からの評価の中にも順次反映されていくと思ひます。

ただし、そもそも学術コミュニティにおいて研究データを管理・利活用していくこと自体が認められないと、本当の意味での定着はしないと思ひています。そういう意味では、このように活用してうまくいったという事例を積み重ねながら、コミュニティの中で、これを評価すべきなのかどうかという議論も並行して進められると、取り組みが全体的に広まっていくのではないかと思ひます。

●**池内** そうですね、評価ありきになってしまっても違うかなという感じはします。先ほど、COVID-19に関してはデータ共有が盛んに進められたという話がありました。バイオの、特にCOVID-19関係のデータ共有に関しては、公開してもそのことによって研究者が不利益を得ない仕組みづくりがされているかと思ひます。先に研究データを公開したとして、例えばそれを使った論文を後から発表しても二重投稿にはしないなど、分野の中でのルールメイキングのようなことも必要なのかなと思ひます。

パネリストの八塚さん、バイオの方は、随分前からデータ共有が当たり前の世界になっていると思ひますが、データ共有をドライブしていくような動きというのは、これまでどのようなものがあつたのでしょうか。

●**八塚** 先ほど大波さんからデータ共有のメリットの

話の中で、COVID-19 の情報共有の例が挙げられて、私はそれを聞いてはっとしました。あまりにも当たり前過ぎて、事例だという意識が全然なかったのです。ということは、恐らくこれからつくられていくであろう研究データポリシーも、あって当たり前で、それが素晴らしく世の中に貢献したといわれて「そうだったっけ？」という世界になっていくのではないかと思います。

私もウイルスは専門ではないので分からないのですが、恐らく COVID-19 の場合も、今までつくられてきた研究コミュニティ内のデータ共有の文化にのっとって研究が行われていると思います。もちろん皆さんスピード感を持って研究していると思いますし、専用のリポジトリが幾つかできているのも知っていますが、そのために大きく仕組みを変えたということではなく、むしろ今までであったものが活用されたという認識をしています。

●池内 バイオの世界ではもはや当たり前ということで、そのうち全分野で研究データポリシーが過去のものになるのかなという感触も受けました。人類未曾有の危機である COVID-19 に関しては取り組みが大変な速さで進んでいて、これはデータ共有の大変良い事例として誰もが腹落ちするような先行事例ではないかと思えます。

会場からの質問がもう少しあるでしょうか。

●林 一つ来ています。「どなたかに質問したら適切か分からないのですが、研究データをオープンにしても、論文がオープンアクセスではない場合、どのように研究が行われたかの詳細が分からず、データの使い道が絞られてしまうのではないかと思います。大手出版社の有料購読アクセスに対して、オープンデータポリシーはどのように対応されるのでしょうか」。なかなか悩ましい質問だと思いますが、いかがでしょうか。山形さん、お願いします。

●山形 私が直答で全部答えられる質問ではないのですが、これまでの図書館の仕事とこれからの仕事をつないでくださるような非常にいいご質問だと思います。現場の図書館員の皆さんの多くは、研究データの話が急にではないですが降ってきて、これまでやったことがないからどうしたらいいのだろうというモチベーションで本日参加されているのではないかと思います。まさにこのようなご質問のところで、私たち図書館員が今まで管理してきた論文とデータが両輪なのだということを強く感じます。諸外国ではゴールドオープンアクセスなども結構強制力を持って進めていて、そういう違いがここで出てきたのだと思います。

論文とセットのデータは、論文とセットで使う限りにはその論文のエビデンスという意味合いが強いのではないかと想像しますが、他方で、データの使い道は、その論文に付随している一つの使い道だけではなく、違う利活用の仕方もあると思います。そういった違う利活用ができるのかどうか、クローズドの購読論文に付随した場合にデータポリシーがどのような効力を持つのかについて、八塚さんや、そのあたりに詳しい方からコメントや事例紹介を頂けるといいのかなと思います。すみません、お答えではなくコメントです。

●八塚 質問の意味を十分に理解できていないのですが、このデータポリシーというのは、ジャーナルの側のデータポリシーという意味でしょうか。

●林 大学からの質問ですから、大学のデータポリシーなのかと私は読んだのですが。

●八塚 なるほど。ジャーナルのデータポリシーが要求しているのは、ジャーナルがある条件を提示した上で認めたリポジトリに必ずデータを置いてくださいということだと思います。(リポジトリに置かれた)データには必ずライセンスが付いているはずですが、そういったライセンスを明示しないリポジトリやデータは、そもそもジャーナルのデータポリシーに反するはずな

ので、そのライセンス次第ということになると思います。結局、利用者にとって意味があるのはライセンスなのです。データポリシーはあくまでも、データを作成した側、データを管理する側がどのようにライセンスを付けていったらいいかという指針であり、利用者にとって意味があるのはライセンスの方です。

●山形 研究の現場にいないのでよく分からないのですが、データに付いているポリシーと論文に付いているポリシーが違ったり、論文に付いているライセンスとそれとセットのデータに付いているライセンスが違ったりすることはあるのですか。

●八塚 あり得ると思います。例えば、論文自体は完全なオープンアクセスになっていないけれども、データはオープンアクセスになっているというものは結構あります。というのは、論文に関するデータを置いているリポジトリは必ずしもそのジャーナルが管理するリポジトリとは限らず、研究コミュニティが運営しているリポジトリだったりするからです。ですから、データのライセンスは論文のライセンスとは全く別物と考えていいと思います。たまたま同じ場合もあるでしょうけれども。

●山形 なるほど。私がもしこの質問をした方だったら、そこの食い違いがあってもいいのか、その食い違いをデータポリシーはどのように扱うのかというのが混乱するところかなと思いました。

●八塚 これは他の方にも伺ってみたいです。私の考えでは、そこはデータポリシーの範囲外ではないかと思います。もちろん異論はあるかもしれませんが、私はそのように認識しています。

●林 他の皆さんはいかがでしょうか。まさに今、対話が繰り広げられているので、乗っかっていただけるとさらによいかと思います。

先ほどの議論ともかぶりますよね。結局、著作権や所有権などを含めてライセンス周りの話に行き着くわけで、それを決める要素には幾つか観点があって、何のルールが優先されるべきかという話に最後は落ち着いていくと思います。このあたりは整理し切れていない部分があって、最終的には法制度の専門家も入れて整理しなければならない話なので、本日は論点出しにとどめさせていただきます。

もう一つ質問をご紹介します。「機関リポジトリ、JPCOAR 等と研究データポリシーとの関係はどのようなものになっていくのでしょうか」。私の解釈では、データポリシーを策定していく上で、JPCOAR としてデータポリシー策定に役立てるような支援があり得るのか、あるいはデータポリシーへのコミットをどうしていくかという方針やビジョンが何かありますかという質問だと思いますが、これはぜひ安原さんにお答えいただければと思います。

●安原 JPCOAR には研究データ作業部会というものが、この研究データ作業部会と AXIES-RDM 部会とが連携し、2020 年度に国内機関における研究データ管理の取り組み状況調査のアンケートを行いました。各参加機関の状況を調査して、集計結果は 2020 年度 RDM 事例形成プロジェクト中間報告書 (<https://jpcoar.repo.nii.ac.jp/records/2000003>) に掲載しています。今年度末には、林先生や池内先生、そして NISTEP の方々に分析協力を頂いて、NISTEP から報告（日本の研究機関における研究データ管理（RDM）の実践状況 <https://doi.org/10.15108/stih.00287>）を頂けるような形です。

また、少しずれますが、JPCOAR も所属する、オープンアクセス支援のための国際連携組織である COAR という団体があります。2021 年 2 月の第 3 回 SPARC Japan セミナーでご紹介しましたが、COAR がつくったリポジトリの評価フレームワークを基に、各機関が自機関の機関リポジトリは国際的に見てどういう機能が足りていないのかということがチェックできるリス

トを JPCOAR で作成しました。今年に入ってそのチェックリストを関係機関の方々に実際に使っていただき、機関リポジトリ/データリポジトリの運用実態に関するアンケート調査を行いました。この調査の結果についても今年度中に中間報告をして、来年度にはさらなる分析結果を報告できると思います。

また、COAR のフレームワークの抽象的な点や分かりづらい点、改善点を、日本から意見を言う形でまとめて COAR の方に報告し、COAR のフレームワークの改善につなげていきたいと考えています。

●池内 今回の SPARC セミナーの狙いもそうなのですけれども、先進的な取り組みをしているところがどのように取り組んでいるか、どこが大変なのかという点は、今から取り組む大学や研究機関にとっては非常に参考になると思います。それを取りまとめて発表してくださることは本当に意義があると思いました。

ここまでのお話を聞いて、電気通信大学の上野さん、まさにデータポリシーの策定に携わっている立場からコメントがありましたらお願いできますか。

●上野 先ほどは少し本学の体制などをお話ししたのですが、実は前段があります。ある先生が論文投稿をする際に、Agora という図書館のセンシングデータを使って研究したのですが、そのデータを公開してほしいという要請が査読の段階からあり、どうしたらいいのだろうという問い合わせがありました。また、学内で研究している中で、COVID-19 に関わるデータを探しているのだけれども図書館の方で見たりしていないかという問い合わせもありました。そのように、先生方がデータを管理したり、オープンになっているデータを探したりする現場にたまたま付き合う機会が何度かあり、それを踏まえて本学としては、本日、田野学長がお話ししたような大方針も踏まえながら、どういう形でポリシーを定めていくのがいいのかということを考えています。

図書館は図書館、先生は先生という乖離がある状態

のままだと、なかなか思うような成果は出ないと思うので、可能な限り、先生方が研究の中でどういう課題に突き当たっているのか、そこに対して図書館は何かできるのかというスタンスで話をし、先生方との関係を持ちながら、それをどうフィードバックしてポリシーをつくるのか、どうシステムを整備するのか、大学のミッションと乖離がないのかということを整理していく必要があると思っています。

●池内 その部分はなかなか図書館が踏み込めなかったというか、しなくてはと思いつつも積極的に関わりづらかった部分かと思いますが、データをきっかけに支援できることが出てきたのかとも考えます。

本日の司会の矢吹さんは研究支援をする URA の立場で大学にお勤めですが、矢吹さんからはいかがでしょうか。

●矢吹 実は私は、この問題の割とそもそも論のところでも毎回つまづくのです。というのは、まさに研究者は政策的に迫られて、やむにやまれずデータをどうにかしなければいけないという状態になっていて、本日の田野学長のお話にもあったように、もっと前向きに取り組んでいくべきということも十分に分かる一方で、フルオープンにしるという誤解があつてややこしくなっていて、「そんなものオープンにできるか」という反発を受けるということは私自身も経験しています。フルオープンにする必要はないと説明するのですが、しばらくするとまた忘れて「全部オープンにしるという意味か」と言われる状況に毎度陥るのです。

政策的にはオープンにするということ自体は前面に出ているはずなのですが、本日の最初の三宅室長のお話を伺って思っていたのは、昨今騒がしい経済安全保障というキーワードも出てくるし、研究インテグリティの問題もかなりクローズアップされている状況でこの研究データの問題に取り組まなければいけないということ。経済安全保障と研究インテグリティの話は、研究者を守るという側面がある一方で、情報を適

切に管理するとき、ある意味ではクローズにするという側面が色濃く出ることがあり、この辺で政策的にかなり齟齬を来しているような気がします。そうなる、データ管理の政策に対して、本気で誰かが最後まで責任を持って進める気であるのだろうかという、言ってしまうと不信が生じる瞬間があると思っています。先生方には「こういう状況になっているので、もっとオープンにした方がいいですよ」と言ってみるものの、やはり無責任に言っているなという後ろめたさも感じています。

これは最終的には各大学や研究者、研究コミュニティの問題なのだ政策サイドからバサッと切られてしまうのではないかと正直思っていますし、実態としては実はそうなのかもしれないという気もします。その辺でどうすればいいのかという悩みを持って本日は勉強させていただきました。

●池内 今回のコメントに対して、三宅室長、いかがでしょうか。

●三宅 元々、研究データ管理においても、当然オープンにしてはいけない情報はあります。例えば個人情報や、海外であれば EU 一般データ保護規則 (GDPR) の関係で非常に管理が難しいものがあります。それ以外にも、例えば権利関係で、特許を取ろうと考えているのであれば最初はオープンにするべきではなく、あるタイミングからオープンにするという情報もあります。また、経済安全保障の話でも、データの管理は当然求められています。ですから、私が強調したいのは、オープンにすることが是ではなくて、しっかりと管理することが大事だということです。

当然、研究の在り方の変容、DX 化という観点であれば、オープンにされたデータをうまく活用していこうという話があります。しかし、その手前の根本のところを考えると、出してはいけないデータはきちんと出さないようにしているということも含めてしっかりと管理することこそが、研究の進捗の上でも意味があ

りますし、説明責任を果たしていく上でも重要です。まずはそういう大前提があった上での話だということをお伝えできればと思っています。

多分、こういった問題は昔からずっと繰り返されているのだと思います。情報技術が発展した中で、特にデータの利活用が生まれたからこそクローズアップされているのですが、実は前からいわれてきたことを改めて整理しているのだと思います。そこをしっかりと具体化することが、研究者の研究能力の向上にもなるし、研究者自身を守ることにもなるということを引きとご理解いただけるよう、説明を尽くしていくことが大事だと思っています。

●池内 研究データの管理という言葉の意味が今のコメントでだいぶ深まったように思います。そして、田野学長が電気通信大学でされていること、しようとしていることは、本当の意味でのデータの管理なのだと思います。ありがとうございました。

話は尽きませんが、お時間となりましたので、総合討論第2部を終了いたします。どうもありがとうございました。