

# 研究データ基盤整備と国際展開

ワーキング・グループ

## 第2フェーズ報告書

令和3年3月

研究データ基盤整備と国際展開ワーキング・グループ



## 目次

○はじめに .....	2
I. 研究データ管理・利活用の考え方 .....	3
(1) データガバナンスの必要性・目的 .....	3
(2) データ収集を目的とする研究に関する考え方 .....	5
(3) 研究者による研究データマネジメントについて .....	6
II. 研究データのあり方 .....	7
(1) 研究データの形式・構造 .....	7
(2) 研究データの質・信頼性とメタデータの標準化 .....	8
(3) 研究データ提供者の真正性 .....	9
(4) プロジェクト終了後などの研究データの取り扱い .....	10
III. 研究データ基盤のあり方 .....	11
(1) 産学で利用可能な研究データ基盤システム .....	11
(2) 効率的な検索機能 .....	11
(3) 産業界の基準を満たすセキュリティ性能と運用体制 .....	13
(4) 持続的な運用の仕組み .....	13
IV. 産学間における研究データ利活用促進のあり方 .....	14
(1) 産学連携の形態 .....	14
(2) 産学で共創するべき研究データの内容の考え方 .....	15
(3) 産学による研究データ共創の事例 .....	16
(4) 産業界による研究データ共有・公開の事例 .....	17
(5) 国による基盤的研究データの整備及び統合の必要性 .....	19
V. さまざまな分野におけるデータの管理・利活用 .....	20
(1) 人文学におけるデータの管理・利活用 .....	20
(2) 農業分野におけるデータの管理・利活用 .....	21
(3) 自動運転における実験データの管理・利活用 .....	22
VI. 高度な研究データマネジメントを実現するための環境整備 .....	23
(1) 研究データマネジメント人材 .....	23
(2) 大学図書館による支援 .....	24
VII. 法的・制度的な課題 .....	25
(1) 研究データの帰属 .....	25
(2) 研究データ授受に関する契約 .....	27
(3) 研究データに個人情報を含む場合の取り扱い .....	27
VIII. 国際連携 .....	29
(1) 国際協調 .....	29

(2) 各国における研究データ基盤との連携 .....	30
(3) 各国際的な枠組み及び組織との連携 .....	30
(4) 国際的共同研究におけるデータ連携 .....	31
IX. まとめ .....	32
(別添1) 本報告書で紹介した事例のまとめ .....	1
(別添2) WG 委員リスト .....	5

○はじめに

研究開発の過程では、実験や観察、計算などによってさまざまなデータが生み出される。またそれらが加工され、プレプリント、論文等の学術的な成果や特許等の知的財産に結びつくまでも多くのデータが生み出される。これらのデータは重要な知的資産であり、これらが積み重ねられ、さらに相互に結びつくことで、新たな科学技術の発展やイノベーションの創出につながる。本報告書では、研究開発の過程で生まれる全てのデータのことを研究データと定義づけるが、科学技術・イノベーションの進展の本質は研究データの蓄積と結合であると言える。

オープンサイエンスは積極的な知の共有により知的生産活動に発展をもたらすものであり、世界的な潮流となっている一方、世界的な出版社等による論文等の研究成果の寡占が進んでいる。また、研究開発の過程で生まれる研究データについても、世界的な出版社やIT企業がビジネスの対象としての関心を持ちつつある。我が国としては、自由でオープンな研究活動を尊重する価値観を共有する国・地域・国際機関等とともに、世界的なオープンサイエンスの発展に貢献することが必要である。また、同時に、我が国の重要な知的資産として、オープン・アンド・クローズ戦略のもとで、研究データを含む研究成果を積極的に利活用していくことが不可欠になっている。現在は、このような移行期の端緒にあるという認識のもと、公的資金による研究データの管理・利活用を進める必要がある。

一方で研究手法においては、現状でも、仮説・検証型を中心とする従来の研究方法が主流であるものの、例えば、マテリアル分野やライフサイエンス分野等においては、データ駆動型科学に代表されるように、知識に基づく仮説を検証していくのではなく、豊富なデータの収集とそれらのデータへの洞察に基づく新たな研究スタイルが見られるようになっている。したがって、研究データの重要性はこれまでと変わるものではないものの、改めて強く認識されてきており、それと同時に、貴重な研究データが散逸しないように適切に管理し、より多くの人々と共有するなど、最大限に利活用することが求められてきている。

科学技術基本法が25年ぶりに改正され、「イノベーションの創出」を基本法の柱の一つに据えるとともに、これまで科学技術の規定から除外されていた「人文・社会科学（法では「人文科学」と記載）のみ」に係るものを法の対象である「科学技術」の範囲に積極的に位置づけた。法改正の背景には、この25年間にイノベーションという概念の含意が大きく変化したことがある。かつて、企業活動における商品開発や生産活動に直結した行為と捉えられがちだったこの概念は、今や、経済や社会の大きな変化を創出する幅広い主体による活動と捉えられ、新たな価値の創造と社会システムそのものの改変を見据えた「トランスフォーマティブ・イノベーション」という概念へと昇華しつつある。元来、人文学や社会科学では、古典籍等の資料を対象として研究を実施するスタイルが多く見られたが、近年では、それら

紙を媒体とした資料のデジタル化とオープン化が徐々に進み、さまざまな研究分野での利活用も進みつつある。

我が国は、研究データのプラットフォーム化の動きや、データ収集力×データ解析力（スパコン・人工知能）が、研究活動に対して圧倒的なインパクトをもたらす可能性に対応する必要がある。

政府としては、第5期基本計画における、「オープンサイエンスの基本姿勢として、公的資金による研究データの利活用を可能な限り拡大する。」という方針のもとで、各種のガイドラインの整備などの施策を行ってきており、次期科学技術・イノベーション基本計画でも、その流れを踏襲し、さらにオープンサイエンスを推進することとされている。

本WGは2019年10月に報告書（フェーズ1）を取りまとめ、政府はこれを踏まえて統合イノベーション戦略2020<sup>1</sup>のもとで、研究データ基盤システムや関係する制度環境の整備、ムーンショット型研究開発制度における先進的なデータマネジメントの推進等を進めつつある。

本報告書では、2019年10月の報告書（フェーズ1）に加え、研究データ基盤の管理・利活用に関する追加的な考え方や方策を示す。

## 1. 研究データ管理・利活用の考え方

### （1）データガバナンスの必要性・目的

研究データを適切に保存し、共有及び公開するといった研究データマネジメントの最も重要な目的は、新たな科学技術の発展やイノベーションの創出である。またこれらに加えて、研究公正への適切な対応も研究データマネジメントの重要な目的となり、研究者自身が主体的に意識して実践すべき要件となりつつある。

特に、より基礎的なフェーズから研究を行う場においては、研究成果が短期的にイノベーションに結びつくことは稀であり、先人たちの研究成果を蓄積し、受け継ぎ、新たな研究成果と結合させることで、大きなイノベーションや、その研究分野の発展につながるものと期待される。したがって、大学等においては、研究データという知的資産が散逸しないように、適切な研究データマネジメントの実施が強く求められる。その一方で、産出される研究デー

---

<sup>1</sup> 『統合イノベーション戦略2020』、内閣府（<https://www8.cao.go.jp/cstp/tougosenryaku/index.html>）

データの規模や種類が急速に拡大していく中、研究者が独自の手法で研究データをマネジメントすることは困難であると考えられる。そのため、研究活動がより大きく進展していくために、研究データマネジメントは重要な要素の一つであり、研究者自身が実施すべきものではあるものの、そのために必要な環境提供やルール作り、また研究者をサポートする仕組みなどは、研究者の所属機関が組織的に取り組むべき事項である。また、研究推進法人等においても、研究者が適切な研究データマネジメントを実施する上での指針を示し、適宜指導を行う等の対応が求められる。

具体的には、大学などの各組織において、研究データマネジメントに関するガバナンスのあり方について定めたデータポリシーの策定をさらに進めるとともに、研究者が研究開始前よりデータマネジメントプラン（DMP）を作成し、適宜修正を加えながら、それに則って研究データマネジメントを実施するなどの体系的な研究活動の再構築が必要となると考えられる。

国際的な取組の中では、こうした改革に対する支援方法の確立に向けた標準化が進んでいる。例えば、Research Data Alliance（RDA）に設置された、有志によるワーキング・グループでは、DMP に書かれた内容に沿って研究データマネジメントを支援する、Machine-actionable DMP と呼ばれるコンセプトのシステムを実現するためのデータモデルが策定されている<sup>2</sup>。このように、ポリシーや DMP の内容を、日々の研究データマネジメント活動に効果的に反映する機能をもつシステムも非常に有効であると考えられる。

政府としては、これらの改革について必要なガイドラインなどを適宜策定することで、適切な研究データマネジメントが実施されるよう働きかける必要があり、これまでも幾つかのガイドライン等を策定、公開している。具体的には、2018年6月に「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」<sup>3</sup>を公開しており、2020年9月時点において、11法人・1機関がデータポリシーを策定、公開済みである<sup>4</sup>。また、2019年3月

---

<sup>2</sup> 『DMP Common Standards WG』、Research Data Alliance（<https://www.rd-alliance.org/groups/dmp-common-standards-wg>）

<sup>3</sup> 『国立研究開発法人におけるデータポリシー策定について』、内閣府（<https://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.html>）

<sup>4</sup> NICT、NIMS、NIED、理研、JAMSTEC、農研機構、国際農林水産業研究センター、森林研究・整備機構、産総研、NIES、JAEA の 11 法人と、JAXA 内の宇宙科学研究所の 1 機関において策定、公開済みである

に「研究データリポジトリ整備・運用ガイドライン」<sup>5</sup>を策定、公開し、国研及び大学等が、国際的にも信頼され、適切に研究データを保存・公開できるリポジトリの整備・運用に関する指針を示した。また、当ワーキング・グループから 2019 年 10 月に第一回目の報告書を公開し、その中で DMP の一般項目についてまとめた<sup>5</sup>。2020 年 9 月時点において、8 省・機関の競争的研究費制度において、DMP の提出を研究実施者に要請する仕組みを導入済みである<sup>6</sup>。

また、大学 ICT 推進協議会 (AXIES) は、2019 年 5 月に「学術機関における研究データ管理に関する提言」<sup>7</sup>を公開しており、ここでは、学術機関において研究データ管理を実施するために検討すべき条件や、実施すべき事項が詳細にまとめられている。また、同協議会からは「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 (案)」<sup>8</sup>が 2020 年 6 月に提示され、年度内に取りまとめられる予定である。

## (2) データ収集を目的とする研究に関する考え方

かつてのヒトゲノムプロジェクトに代表されるように、大規模なデータを収集することを目的とした研究は、概して長期的なものになる。また、このような研究手法は、得られた研究データそのものだけを成果と捉えるのではなく、むしろ、次世代の研究者などによってそれら研究データが利活用され、その結果、新たな研究成果やイノベーションが生み出される期待のほうが大きという特徴がある。したがって、特に、データ収集を伴う研究については、長期的な視野に基づいて投資をするといった考え方が必要である。

また、将来における利活用なども見据えた研究のマネジメントや評価を導入することや、収集された研究データの適切な管理や利活用を実行するための支援人材、さらにそれら研究データのための十分なストレージの確保などが今後必要になると考えられる。

---

<sup>5</sup> 『国際的動向を踏まえたオープンサイエンスの推進に関する検討会』、内閣府 (<https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/index.html>)

<sup>6</sup> 文部科学省、経済産業省、AMED、JST、JSPS、NEDO、BRAIN、環境再生保全機構において導入済みである。

<sup>7</sup> 『学術機関における研究データ管理に関する提言』、大学 ICT 推進協議会 (<https://axies.jp/report/publications/proposal/>)

<sup>8</sup> 『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』、大学 ICT 推進協議会 (<https://www.nii.ac.jp/openforum/upload/feb0fce70685554c4d47170d4c189fc5eacde1f6.pdf>)



### (3) 研究者による研究データマネジメントについて

一般的に研究データは、時に研究者自身の独自の思考の源泉となり、時にそのような独自の思考の正当性を判断するための重要な材料となる。したがって、研究分野にもよるが、研究者は通常、自身の研究の独創性が他の研究者によって模倣されないために、取得した研究データは論文などの形で発表するまでは非公表にしたいと考える傾向がある。これは、個々の研究者が互いの研究の独創性や先進性に基づいて、それぞれの研究成果の科学的価値や経済的価値を競う、競争ベースの研究パラダイムにおいては、適切な行動である。

したがって、研究者がこれからの協調ベースの研究パラダイムに向けて、研究データの共有や公開をするようになるためには、相応のインセンティブを設計する必要がある。具体的には、アカデミアや産業界が共同して取り組む大型研究プログラムにおいて、参加主体が互いの研究データを適切に共有できるようにし、研究者、プログラム、機関等の評価基準に関して、従来の論文や学会発表による基準に加えて、研究データの共有や公開に基づく基準を採用することなどが考えられる。

具体的には、単に多くの研究データを共有あるいは公開すればよいとするのではなく、後述するように、各研究者が自身の研究データの共有や公開において、適切なオープン・アンド・クローズ戦略<sup>9</sup>を策定しているかといった研究データマネジメントの戦略としての妥当性や、その戦略に基づいて研究データマネジメントを適切に実施しているかといった実効性などが基準になると考えられる。また、COVID-19に関する研究成果のように、特に速報性が重視される研究成果を発表する場合や、査読論文誌への投稿前にプライオリティを確保するためや議論を誘起するために発表する場合において、プレプリントでの発表を行うことを促進し、以て、データの共有公開も促進することも考えられる。

このような研究者による研究データマネジメントに向けて、まずは、研究者が共有したいと思う研究データの共有や公開から始め、新たな研究パラダイムへの移行とともに、共有及び公開を対象とする研究データの範囲を拡大していくことが肝要である。

またそもそも、特に公的資金が投入された研究データについては、適切に管理され、広く国民に還元されるべきものであるという考え方について、研究者の理解を得ることが重要である。このためには、研究データ管理には相応の時間と労力が必要なことを念頭に、そのための経費や研究支援体制が整備される必要がある

さらに、安定的な研究データの管理のためには、経済的な対価をインセンティブとするこ

---

<sup>9</sup> ここではデータの特性から公開すべきもの（オープン）と保護するもの（クローズ）を分別して公開する戦略を意味する。

とも考えられる。すなわち、公開された研究データが他の研究者などによって利活用される際に、研究データを公開した研究者に利用料が支払われる仕組みなどが考えられる。

。

ただし、ここで留意すべき点としては、研究データの公開、共有及び非公開にあたっては、戦略的な対応を要するという点である。上述のとおり、研究データは研究者の独創的な思考やアイデアと強く紐づくものであるため、各研究者がやみくもに研究データを公開することは、逆に我が国の研究開発や産業の競争力を損なわせることにつながる可能性がある。したがって、各研究者はどの研究データを共有あるいは公開すべきかを適切に検討し、自身の研究成果を最大化しながら将来のイノベーション創出に大きく寄与できるようなオープン・アンド・クローズ戦略に基づいて、研究データの管理を実施することが重要である。

また日本学術会議では、研究データの重要性が著しく高まっている中、研究者が負担を感じることなく研究データを適切に取り扱うことができるよう、オープンサイエンスの深化と推進に関する検討委員会において、2020年7月に「オープンサイエンスの深化と推進に向けて」<sup>10</sup>と題した提言をまとめた。同提言では、①データが中心的役割を果たす時代のルール作りの必要性、②データプラットフォームの構築・普及の必要性、③第1次試料・資料の永久保存の必要性に関する提言がまとめられており、政府としては、これを踏まえて対応することが望ましい。

## II. 研究データのあり方

### (1) 研究データの形式・構造

研究データが適切な共有・公開のマネジメントを通じて、幅広く利活用されるためには、各データの形式及び構造が標準化されていることが必要となる。研究分野によっては、研究データの形式と構造について国際的な標準が定められているケースもある。例えば、太陽地球系科学分野における研究データ（自然現象の時系列データやシミュレーションデータなど）については、NASAが主導した国際的な標準様式がある。その一方で、国際的な標準等が定められていない研究分野においては、研究者が使用する測定機器の種類によって研究データの形式や構造が異なり、研究データの利活用を促進する上で妨げとなっている場合がある。

---

<sup>10</sup> 『オープンサイエンスの深化と推進に向けて』、日本学術会議 (<http://www.scj.go.jp/ja/info/kohyo/kohyo-24-t291-1-abstract.html>)

この課題を解決するためには、各測定機器メーカーとも連携して、業界ごとあるいは分野ごとの研究データの標準化を図ることが効果的であると考えられる。このような取組は、これからの我が国におけるデータ駆動型研究を促進する上でも重要な取組となると考えられ、政府や学会としても本取組について早急に検討を進める必要がある。

その際に留意すべき点としては、測定機器メーカーによっては、単に測定機器を販売するだけでなく、そこから得られる独自の形式・構造のデータに応じた解析ソフト等も販売するというビジネスモデルを構築している点である。したがって、このような測定機器メーカーのビジネスを大きく棄損することのないようにデータ形・構造の標準化の取組を進めていくことが重要である。

また、標準等が定められていない場合におけるその他の対応策としては、データ形式・構造を変換する機能を利用する方法が挙げられる。例えば、農業データ連携基盤（WAGRI）<sup>11</sup>は、データ連携基盤への参加者（農機メーカー等）からの需要が大きい外部データについて、その形式・構造をデータ連携基盤が採用する形式・構造に変換した上で取得可能としている。

## （２）研究データの質・信頼性とメタデータの標準化

2019年末から新型コロナウイルス感染症（COVID-19）が急速に世界中へと拡大していく中、これを可及的速やかに終息させるため、各医療機関や研究機関、研究支援機関、論文出版社等が、新型コロナウイルスに関連する研究成果とデータを広く迅速に共有する声明に署名し、COVID-19に関する論文のオープンアクセス化や、プレプリントサーバー等での迅速な公開、研究データの国際的な共有が行われた<sup>12</sup>。そのような中、国際的医学誌であるThe Lancet及びThe New England Journal of Medicineに掲載されたCOVID-19関連の論文が相次いで撤回された。その理由は、いずれの論文も、データを扱った会社が公開を拒否した不透明なデータを用いて作成されており、データの信ぴょう性が保たれないためであった。このことは、客観的なエビデンスとなる研究データを検証せずに論旨の妥当性とインパクトを判断し、その結果、誤った情報を掲載する論文誌が存在するという重大な問題点を明らかにした。COVID-19に関する論文ということで速やかな公開が強く求められていたという背景があったにせよ、改めて、研究データ公開の必要性と、それを踏まえた上での査読プロ

---

<sup>11</sup> 『農業データ連携基盤(WAGRI)の本格運用について』、農研機構 ([https://www.naro.affrc.go.jp/project/research\\_activities/laboratory/rcait/121064.html](https://www.naro.affrc.go.jp/project/research_activities/laboratory/rcait/121064.html))

<sup>12</sup> 『Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak』、Wellcome (<https://wellcome.org/coronavirus-covid-19/open-data>)

セスの重要性が示された。

研究データの利活用を促進する上では、各研究データの信頼性を確かめる手段が不可欠である。従来であれば研究者は、論文発表や学会発表などのプロセスを通じて、データ取得時の実験や観測、計算等の条件（用いた手法や材料、測定装置、測定条件等）を詳細に示すことで、自身の研究データの質や信頼性を説明してきた。上記の論文撤回の事例においても、扱われた研究データがどのようなものであるかについて論文中で説明されたことで、その内容が従来の知見と矛盾する箇所があったことが明らかになり、信頼性に疑念が持たれるきっかけとなった。

一方、このような論文発表や学会発表といったプロセスを経ずに、研究データのみが利活用される場合において、その質や信頼性を示す手段としてはメタデータが挙げられる。メタデータによって、データの質や信頼性を示す取組事例として、一般社団法人データ流通推進協議会はデータカタログ作成ガイドラインの策定を進めている<sup>13</sup>。同ガイドラインは、組織や分野を横断したデータの検索や比較を実現するため、共通形式でデータの概要を表現できる項目定義を目指している。

また、メタデータは研究データの質や信頼性を示す手段として有用なだけではない。後述するように（Ⅲ. 研究データ基盤のあり方（2）効率的な検索機能）、分散管理された研究データの中から、所望の研究データを検索する上でも、メタデータを活用することは非常に有効な手法であると考えられる。

### （3）研究データ提供者の真正性

研究データの利活用を促進する上では、各研究データ提供者の真正性を確かめる手段が不可欠である。具体的に研究データ提供者の真正性を確かめるには、研究データ提供者を、個人主体の方法で識別するのではなく、所属組織との関係の中で識別することが重要である。

所属組織との関係の中で個人を識別するという方法を取れば、個人主体の方法で識別する場合と異なり、研究データ提供者の属性情報（どの研究所に属する研究者なのか等）まで確認される。その結果、研究データ提供者の真正性だけでなく信頼性も確認することができる。また、研究データの信ぴょう性や質における責任は、研究データ提供者だけでなく、その所属組織も負うことになるため、研究データそのものの信頼性にもつながる。

---

<sup>13</sup> 『公開資料』、一般社団法人データ流通推進協議会（<https://data-trading.org/public-release/>）

#### (4) プロジェクト終了後などの研究データの取り扱い

公的資金が活用されたプロジェクトの場合、プロジェクト終了後において研究データが適切に管理されず、散逸してしまうケースがある。このような事態を防ぐためには、プロジェクト終了後も見据えた、研究データの保存・管理のための人的および経費措置が求められるとともに、プロジェクト開始前から、管理すべき研究データに関する DMP を策定し、適宜必要に応じて改定しながら、適切なデータマネジメントを実施することが必要となる。

例えば、研究データを大学の機関リポジトリにおいて管理・公開する場合、研究プロジェクト経費から機関リポジトリ維持費に支出できる仕組みなどについても検討されるべきである。あるいは、その維持は大学の大きな責務であり、間接経費などの一般経費の活用も一つの選択となりうる。

また、プロジェクトの中間評価や最終評価などの時間的な節目において、保存すべき研究データの見直しや、保存すべき場所などについての検討を行うための仕組みを導入することも重要である。また、研究者の異動や退職を想定し、予め研究者と所属機関の責任関係の明確化や、外部からの競争的研究費により得られた研究データの取扱い等について、明確にしておく必要がある。

例えば、京都大学では、各組織のレイヤーに応じて研究データの必要性について精査し、必要と判断した組織がその管理費用を負担する仕組みを検討している。具体的には、大学が残しておくべきと判断した研究データの管理費用については大学が負担するが、大学が不要と判断した研究データについては、各部局でその必要性について検討し、その結果、保存しておくべきと判断した場合については、部局がその管理費用を負担するというものである。

以上のように、時間や組織などの観点で、それぞれの研究データの必要性を適切に判断する仕組みを導入することは、有用な研究データを効率よく蓄積していく上で効果的であると考えられる。ただし、このような仕組みを円滑に実行していくには、研究者自身がデータマネジメントをしやすいように研究データを作成し、又は、いかに低コストで研究データを保管できるようにするかなどの手法について、大学院生や若手研究者の時期から適切な研修等を通じて身につけることが重要である。

一方で、より多くの有用な研究データを蓄積していくことが、我が国のこれからの研究開発のさらなる発展やイノベーションの実現において必須であることは間違いないものの、爆発的に研究データが生み出されていく中で、それらを保管するのに十分な容量のストレージを確保しておくことは将来的に重要な課題となる可能性がある。研究者が不自由なく

研究データを生み出す環境を維持し続けるためにも、どのような仕組みでこの課題を解決すべきか、どのように必要な費用を分担すべきか等、今後検討する必要がある。

### III. 研究データ基盤のあり方

#### (1) 産学で利用可能な研究データ基盤システム

大学・研究機関が公的資金を活用して生み出した研究データを適切に保存、共有、公開し、民間企業がそれらを利活用して新たなイノベーションの創出へとつなげる活動を促進するためには、アカデミアと産業界が共同で利用できる研究データ基盤システム（研究データの保存、共有及び公開が可能なシステム）が必要であるが、現状、我が国や海外において整備が進められている研究データ基盤システムはいずれもアカデミアを対象としている。

例えば、世界的にも最も開発が進んでいる研究データ基盤システムである European Open Science Cloud (EOSC)<sup>14</sup> についても、また、我が国の中核的な研究データ基盤システムとして国立情報学研究所が開発し、2020年度中に本格運用を開始する NII Research Data Cloud<sup>15</sup> についても、現時点では、主としてアカデミアを対象としたものである。しかし、EOSC に関しては、欧州における産業界のデータマネジメントのためのフレームワークとして、ドイツとフランスが主導して立ち上げた GAIA-X<sup>16</sup> と連携する動きも見られる。また、NII Research Data Cloud に関しては、アカデミアのみならず産業界も円滑に利用できるようになる予定であり、産学の研究機関が参画するムーンショット型研究開発制度においても、先行的に同システムの活用を図るなど、先進的データマネジメントを推進することとされている。  
エラー! ブックマークが定義されていません。

#### (2) 効率的な検索機能

実際の研究データは、さまざまな分野別リポジトリや各機関の機関リポジトリなどにおいて分散的に保存されているのが一般的である。このような状況において、所望の研究デー

---

<sup>14</sup> 『European Open Science Cloud (EOSC)』、European Commission (<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>)

<sup>15</sup> 『NII 研究データ基盤 (NII Research Data Cloud) の概要』、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 (<https://rcos.nii.ac.jp/service/>)

<sup>16</sup> 『GAIA-X』 (<https://www.data-infrastructure.eu/GAIA/Navigation/EN/Home/home.html>)

タを利活用したい産業界の研究者などにとっては、分散保存された研究データの内容や保存場所等が記載されたメタデータを利用することが非常に有効であると考えられる。

例えば、経済産業省では、所管するナショナルプロジェクトにおいて、研究開発データの利活用促進のため、各データのメタデータをデータカタログ化して、HP上で公開する取組を開始している<sup>17</sup>。ここでのメタデータ項目は、提出日、データ名称、データの説明、管理者、プロジェクト終了後のリポジトリ、概略データ量、研究データの利活用・提供方針、連絡先、契約件名（プロジェクト名）、備考であり、経済産業省のHPで一般に閲覧することができる。

また、国立極地研究所や名古屋大学宇宙地球環境研究所など、国内の7大学・研究所から組織される大学間連携プロジェクトである「IUGONET」<sup>18</sup>では、分散管理されている太陽・地球物理に関する研究データの異分野間での流通を進め、新たな課題研究を推進することを目的として、メタデータのデータベースをネットワーク上で共有し、異分野の研究者でも容易に各研究データの概要を知ることができ、かつ、それら研究データへアクセスできるシステムを構築している。その結果、融合科学への発展が確立されつつある。

メタデータを活用する手法は、研究者が、自身の研究データを一般に公開したくはないが、所望の条件を満たす第三者となら共有してもよいと考えている場合において特に有効であると考えられる。すなわち、実際の研究データの情報のうち、公開可能な情報についてのみをメタデータとして公開し、これを見て興味を持ち、連絡してきた研究者等がいれば、当事者間で交渉をして実際の研究データを共有するか否かの判断を行えばよい。

このような仕組みがうまく機能すれば、産学における特定の関係者間でのデータ共有やそこからのイノベーション創出を促進できると期待される。そしてその実現には、研究データを利用する研究者が、メタデータの情報を介して目的の研究データに速やかにたどり着くことができるために、メタデータを効率的に検索可能な仕組みを構築することが重要である。

そこで内閣府では、上記の事例に倣って、研究データ基盤システム(NII Research Data Cloud)を用いることで、公的資金が投入された研究データのメタデータを効率的に検索できると

---

<sup>17</sup> 『委託研究開発におけるデータマネジメントに関する運用ガイドラインとナショナルプロジェクトカタログ』、経済産業省 ([https://www.meti.go.jp/policy/innovation\\_policy/datamanagement.html](https://www.meti.go.jp/policy/innovation_policy/datamanagement.html))

<sup>18</sup> 『IUGONET Web Service Type-A』、IUGONET (<http://search.iugonet.org/list.jsp>)

ともに、それらの研究データへ容易にアクセスできる仕組みを構築することを目指している。NII Research Data Cloud では、さまざまなリポジトリに保存された研究データのメタデータを集約し、同システムの検索基盤（CiNii Research<sup>19</sup>）を通じて、誰でも検索可能な仕組みを実現する予定である。将来的には、公的資金による研究データの積極的な利活用を推進するため、研究データにメタデータを付与し、全てのメタデータを同検索基盤によって検索可能となることが望ましいと考えられる。

このことと関連して、内閣府等の関係府省が推進するムーンショット型研究開発制度<sup>20</sup>において、NII Research Data Cloud を用いるなどによって、研究推進法人が、PM 及び研究者に提出させたメタデータから構成されるデータカタログを管理する取組を進めている<sup>21</sup>。

### （3）産業界の基準を満たすセキュリティ性能と運用体制

上記のように、産学の両者によって共通で同じ研究データ基盤システムが利用される場合において、重要な課題の一つがセキュリティの問題である。一般的に産業界のセキュリティの基準が厳しいため、その基準を満たすものでなければいけない。また、単にストレージやシステムだけでなく、運用体制も含めて、十分なセキュリティ体制が構築される必要がある。セキュリティの対象となる情報としては、研究データそのものに関わる情報もちろんのこと、例えば、検索履歴なども各企業の研究開発戦略に紐づく重要な情報となり得るため、十分に守られる必要がある。

### （4）持続的な運用の仕組み

研究データ基盤システムを用いた産学の研究データ管理・利活用に関する活動を継続的なものにするためには、研究データ基盤システムの運用等が、将来的には公的資金だけに依存するのではなく、システム利用者等のステークホルダーからの資金によって賄われる、自律性のある仕組みとなることが理想である。その具体的な方策の一つとしては、研究データ基盤システムを介して研究データを利活用する際に、それに相応しい利用料を負担する仕組みも考えられる。

---

<sup>19</sup> 『CiNii Research (検索基盤)』、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 (<https://rcos.nii.ac.jp/service/research/>)

<sup>20</sup> 『ムーンショット型研究開発制度』、内閣府 (<https://www8.cao.go.jp/cstp/moonshot/index.html>)

<sup>21</sup> 『ムーンショット型研究開発制度の運用・評価指針』、内閣府 (<https://www8.cao.go.jp/cstp/moonshot/shishin.html>)



より具体的には、製薬や機械、電気、鉄鋼、化学などのさまざまな事業分野において、データ駆動型の研究開発活動の重要性が増す中、各民間企業単独では十分なデータ量を生み出すことは極めて困難になってきているという背景がある。したがってアカデミア等が、民間企業にとって利用価値の高い研究データを生み出し、適切に共有あるいは公開すれば、各民間企業による研究データの利用料を運営費に活用することが考えられる。このような仕組みを通じて、国益としての研究データを継続的に管理し続けるという考え方が重要である。

#### IV. 産学間における研究データ利活用促進のあり方

##### (1) 産学連携の形態

アカデミアが生み出した研究データを民間企業が利活用する場合や、アカデミアと民間企業が互いに研究データの共有や公開しあう場合、またはアカデミアと民間企業が共同で新たな研究データを創出する場合などにおいては、両社の間での何らかの合意が必要となる。このような場合、互いの秘密情報を第三者に漏洩しないために、両者間で信頼関係があることが重要である。

産学間での連携の形態としては、1対1や1対複数、複数対複数のケースが考えられるが、信頼関係を構築しやすいという観点を考慮すれば、関係機関をできるだけ少なくした1対1の連携が適していると考えられる。このような場合、関係機関間で最適な契約を取り交わしやすく、互いの情報交換においても第三者への漏洩のリスクを極めて低く抑えることが可能なため、安心感を持って積極的に情報交換を行いやすい。しかしその一方で、1対1での連携においては、限られた機関間での研究データ利活用しか実施されない。したがって、ある特定の民間企業が複数のアカデミアの研究データを利活用したい場合においては、対象のアカデミアの機関の数だけ契約を交わす必要があり、多くの労力が必要となる。また、データ利活用に関するルールもそれぞれの連携ごとにより変わるため、実務上においても煩雑になる。

一方で、コンソーシアムのような複数の産学の機関間での連携においては、複数の民間企業に研究データの利活用の機会が与えられる。このことは、公的資金によって生み出された研究データの価値を最大化するという観点から利点があると考えられる。また、その他のメリットとして、ルールの統一化、手続きの簡略化・標準化が挙げられる。これにより、各民間企業はそれぞれ個別の権利関係に煩わされることなく、複数のアカデミアの研究データを利活用できることが期待される。しかしながらその一方で、参画機関が多くなればなるほ

ど、各機関間での信頼関係の構築は困難になる。このことにより、各機関は必要な情報開示をためらい、十分な研究開発を達成できなくなる可能性もある。

以上のように、産学間の連携のあり方においては、その手段において一長一短があることを留意した上で制度設計をする必要がある。例えば、一般社団法人産業競争力懇談会(COCN)のプロジェクトの「エネルギー革新に向けた MI 基盤の構築」<sup>22</sup>においては、それぞれの産学連携の特徴をうまく組み合わせた次のような仕組みを提言している。すなわち同プロジェクトにおける提言では、機能性高分子材料のデータ駆動型科学による開発加速という目標において、まず、拠点となる大学と企業が1対1で連携し、信頼性を担保した上でスモールスタートを実施することが重要であるとしている。そして、1対1の産学連携の中で、互いのデータベースによるデータ連携の有用性を確認した後、大学を核としたコンソーシアムを形成し、各社の協調領域を見極めた上で、大学のデータベースやインフラの共用化、また、新たなデータ取得のための材料開発を実施するとしている。このように、研究開発の各フェーズなどに応じて産学連携の実施形態を工夫することで、それぞれのフェーズに適したデータマネジメントを実施することが可能となると考えられる。

## (2) 産学で共創するべき研究データの内容の考え方

アカデミアが生み出す研究データが民間企業によって最大限に利活用されるためには、それら研究データの内容が民間企業のニーズと合致している必要がある。しかし、アカデミアだけでその判断をするのは難しい。したがって、公的資金による産学の共同プロジェクトなどにおいては、プロジェクト開始前や初期の段階により、予め民間企業とアカデミアが社会実装やイノベーションの実現において必要と考えられる研究データの内容や、その研究データを得るための技術的困難性などについて十分議論し、合意するための仕組みなどを導入することが重要であると考えられる。

その際に留意しておかなければならない点は、それぞれの産業界におけるオープン・アンド・クローズ戦略である。各民間企業における研究データや保有リソース(開発中の素材など)の協調領域と競争領域を考慮に入れ、その共同プロジェクトにおいて産出するべき研究データを定めることが求められる。

---

<sup>22</sup> 『COCN フォーラム 2019 エネルギー革新に向けた MI 基盤の構築』、一般社団法人産業競争力懇談会 (<http://www.cocn.jp/forum/2019/>)

### (3) 産学による研究データ共創の事例

必要とされる研究データの内容や、産業界におけるオープン・アンド・クローズ戦略について、予め丁寧に検討した上で、産学共同での研究データ創出を推進している事例として、NIMSと三菱ケミカル(株)、住友化学(株)、旭化成(株)、三井化学(株)による、NIMSを中核とした化学業界におけるオープンプラットフォーム(化学MOP: Materials Open Platform)の取組がある<sup>23</sup>。同化学MOPにおいては、化学メーカー各社が自社の保有する高分子材料をNIMSに提供し、NIMSはそれらについて統一された条件のもと、計測データを収集し、マテリアルズ・インフォマティクス(MI)などを用いて評価する。その結果、得られた基礎的な研究データは各社へと共有され、新たな材料開発等へと展開される。

NIMSは同化学MOPを実行する上で、化学業界における各企業との間で多くの議論を積み上げ、取り組むべきテーマの選定を行った。そして、各社との協調領域(高分子材料の構造・物性などに関するデータ集積、MIを活用した情報解析手法の適用と評価)と競争領域(協調領域で得られた研究成果を基に二者間での連携へと発展)を明確にした上で、同化学MOPを推進している。また、NIMSは鉄鋼メーカー三社(日本製鉄(株)、JFEスチール(株)、(株)神戸製鋼所)とも連携して、鉄鋼業界におけるオープンプラットフォーム(鉄鋼MOP)の取組を推進している<sup>24</sup>。

また、アステラス製薬(株)、田辺三菱製薬(株)、第一三共(株)の製薬メーカー三社は、民間企業主導型の新薬探索プログラム「JOINUS (Joint Open Innovation of drug repositioning)」<sup>25</sup>を推進している。同プログラムでは、上記三社が、臨床試験を実施した後に開発を中止した化合物、あるいは前臨床試験を実施した後に開発を中止した化合物を集めたドラッグ・リポジショニング化合物ライブラリーを作成し、公募によって選ばれた国内研究機関(大学、公的研究機関、企業等)の研究者が、独自の評価系でそれらライブラリーの化合物を評価する。そして、そこで得られた研究データは各社へとフィードバックされる。これによって、ライブラリー化合物の新たな適応症を発掘し、新薬開発につながる事が期待される。

---

<sup>23</sup> 『物質・材料研究機構(NIMS)と化学4社によるオープンイノベーションを推進する枠組みの構築』、NIMS (<https://www.nims.go.jp/news/press/2017/06/201706190.html>)

<sup>24</sup> 『物質・材料研究機構(NIMS)と鉄鋼3社によるオープンイノベーションを推進する枠組みの構築』、NIMS (<https://www.nims.go.jp/news/press/2017/06/201706300.html?rss>)

<sup>25</sup> 『新薬探索プログラム<sup>®</sup>』、田辺三菱製薬(株)、第一三共(株)、アステラス製薬(株) (<https://open-innovation-joinus.jp/joinus/>)

また、武田薬品工業（株）（以下、武田薬品工業）は、一般住民の全ゲノムリファレンスパネルの充実と日本人集団に特徴的な遺伝子型の発見、及び全ゲノム情報と脳MRI画像を含む健康情報・医療情報の統合的解析による新薬・治療法の研究開発を目的として、東北大学東北メディカル・メガバンク（ToMMo）と共同研究を推進している<sup>26</sup>。同共同研究において武田薬品工業は、資金提供を行う代わりに、本共同研究を通じて取得される約1万人の全ゲノム情報への1年間の優先的アクセス権を獲得する。その後、武田薬品工業とToMMoは、我が国のゲノム情報の医療分野での活用を推進するため、約10万人規模の全ゲノム解析とこの解析結果の利活用を目的としたコンソーシアムの実現を目指している。

以上の事例はいずれも、産学間において、産業界が必要としている研究データの内容を共有した上で、それらデータを共創している事例と言える。また特に、武田薬品工業とToMMoの共同研究の事例は、特定の民間企業が優先的アクセス権というインセンティブを得ることで、将来的にその他の民間企業へも広く展開される研究データをアカデミアと共創するという好事例になると期待される。また同事例は、上述のCOCNプロジェクト「エネルギー革新に向けたMI基盤の構築」における提言と同じように、まずは限定された産学連携でスモールスタートを行い、その後、幅広く研究データを共有あるいは公開していく事例と捉えることができる。

参画企業に優先的アクセス権を与えることで、最終的に広く共有あるいは公開される研究データを産学が共創する事例は海外にも見られる。例えば、米国のバイオテクノロジー企業であるリジェネロン社は、45万人の英国バイオバンク参加者のゲノム中のエキソン配列を決定することを目的として、バイオ医薬品会社のコンソーシアムを主導しているが、このプロジェクトにおいても、参画企業が優先的に研究データにアクセスできる期間が設定されている<sup>27</sup>。

#### （４）産業界による研究データ共有・公開の事例

産学間における研究データの管理・利活用の活動がより発展的に拡大していくためには、共有あるいは公開可能な研究データの量の増大と種類の多様化が欠かせない。その意味に

---

<sup>26</sup> 『2020年 プレスリリース・研究成果 我が国のゲノム医療の大規模な基盤構築へ～武田薬品とToMMoとの共同研究について～』、東北大学 (<https://www.tohoku.ac.jp/japanese/2020/04/press20200414-03-tommo.html>)

<sup>27</sup> 『Exome sequence data available on 50,000』、Biobank (<https://www.ukbiobank.ac.uk/2019/03/new-data-available-exome-sequence-data-on-50000-participants/>)

においては、アカデミア独自の研究データの共有・公開や、前項で示したようなアカデミアと民間企業による研究データの創出だけではなく、民間企業が、少なくとも公的資金を活用して生み出した研究データについては、適切に保存されることは当然のこと、第三者への共有や公開が最大限に行われることが重要である。一般的に民間企業にとって、競争力維持の観点から、自社の研究データを共有、公開することは難しいと考えられるが、最近になって、研究開発フェーズによっては民間企業が自社の研究データを共有あるいは公開する事例が出てきている。

例えば製薬業界においては、かつては臨床試験で得られた被験者レベルの詳細な個別被験者データ（Individual Participant/Patient Data：IPD）を、製薬企業自らが積極的に公開し、一般の研究者と広く共有することは稀であった。しかし近年、臨床試験データは高い透明性が求められるべきという考えと、臨床試験データは公共の財産であり、公衆衛生の向上をもたらす研究に積極的に活用すべきとの考えに基づいて、主として、欧米の製薬企業により IPD の公開・共有が行われるようになった。さらに、2016 年 1 月に医学雑誌編集者国際委員会（ICMJE）が、IPD の公開・共有を医学論文の投稿条件として提案したことにより、我が国においても、臨床試験データの共有（Clinical Trial Data Sharing：CTDS）という言葉が広く知られるようになり、活発に議論が行われるようになってきた<sup>28</sup>。

また同じく製薬業界において、国立研究開発法人日本医療研究開発機構（Japan Agency for Medical Research and Development：AMED）が推進する「産学連携による次世代創薬 AI 開発」<sup>29</sup>プロジェクトでは、産学の保有する生体分子群と化合物群の親和性データや企業の化学研究者が持つ構造最適化に関する経験知データ等、創薬研究における多面的で膨大なデータを広く集約することで、産学が利用可能な化合物設計 AI を開発することを目的としている。同プロジェクトでは、federated learning 技術を活用することで秘匿のままの研究データ提供を可能にしており、これにより民間企業も自社の研究データを提供しやすい仕組みを構築している。開発された AI はその開発に貢献した参加企業に提供することで企業内での研究開発に活用される予定である。

AMED と同様のプロジェクトは海外でも進められている。欧州の製薬企業 10 社とアカ

---

<sup>28</sup> 『臨床試験の個別被験者データの共有 CTDS (Clinical Trial Data Sharing)』、日本製薬工業協会 (<http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/ctds.pdf>)

<sup>29</sup> 『産学連携による次世代創薬 AI 開発』、AMED ([https://www.amed.go.jp/program/list/11/02/001\\_02-04.html](https://www.amed.go.jp/program/list/11/02/001_02-04.html))

デミア、IT 企業が参画した「MELLODDY (Machine Learning Ledger Orchestration for Drug Discovery)」<sup>30</sup> プロジェクトにおいても、federated learning 技術が用いられている。これにより、各製薬企業が自社の研究データを秘匿のまま提供することができ、それらのデータを基に、正確かつ効率的な創薬を実現するための AI 技術の開発を推進している。

以上のように、製薬業界における各民間企業が自社独自の化合物や研究データを提供しあうプロジェクトが生まれている。その背景には、製薬業界においてデータ駆動型の研究開発が進み、より多くの、かつより多様な研究データが必要となってくる中で、それを一社ではとても賅えないという事情がある。同様の変化は製薬業界だけでなく、その他の材料分野でも起こってきており、今後それらの分野においても民間企業による研究データの提供及び共有化が必要になってくると考えられる。

したがって政府としては、そのような研究開発を促進するために、各民間企業におけるオープン・アンド・クローズ戦略をよく踏まえ、民間企業が研究データを提供するための適切なインセンティブを設計したプロジェクト等を推進することが重要となる。

また、民間企業による研究データの提供が増えていった場合には、それら研究データをアカデミアが利用する事例も増えていくと期待される。そのような場合、上述の、研究データ利用者が利用料を支払う運用モデルにおいては、アカデミアが利用料を支払うような相互性を確保する必要がある。なお、一般には、アカデミアの利活用については、いわゆるアカデミック・ディスカウントの対象となる場合が多く見られることを付言する。

#### (5) 国による基盤的研究データの整備及び統合の必要性

産学による研究データの利活用に関する活動を拡大させていく上において、国に求められる役割も大きい。例えば、国民のゲノムデータなど、その分野の基盤となるような膨大な研究データの収集については、国が主導してプロジェクトを推進していく必要がある。また、取得された多様な研究データを連携、あるいは統合し、その価値をより高めることも重要である。

---

<sup>30</sup> 『Machine Learning Ledger Orchestration for Drug Discovery』、MELLODDY (<http://www.melloddy.eu/>)

## V. さまざまな分野におけるデータの管理・利活用

### (1) 人文学におけるデータの管理・利活用

人文学の分野においては、古典籍などの過去の資料のデジタル化とオープン化が徐々に進みつつあり、今後もその動きを継続的に後押しする必要がある。例えば、ROIS-DS 人文学オープンデータ共同利用センターでは、人間文化研究機構・国文学研究資料館がデジタル化した日本古典籍の画像を日本古典籍データセットとして公開するとともに、これから派生したくずし字データセットや江戸料理レシピデータセット、顔コレデータセットなども公開し、さまざまな分野に利活用を広げている。

一方、論文や書籍については、紙の論文や書籍を出版することが業績として認められた時代が長く、そのことを前提とした学術情報流通システムが強固なため、成果物のデジタル化への取り組みは大きく遅れている。そして、研究プロセスを含めた研究データについても、伝統的な出版以外の方法で共有・検証・再利用することへのインセンティブは低いままである。さらに、研究者個人が築いた人的ネットワークから得られた研究データ等については、高い所有意識や調査対象者・地域への配慮などから共有や公開が進まない傾向にある。その結果、他の研究者が同じような研究データを繰り返し収集せざるを得ないなどの弊害が生じる。このような課題を解決するためにも、また、研究成果の公正性を担保する上でも、研究データの共有・公開が図られる仕組みを作る必要がある。このように、先人の成果にアクセスしづらく、データの利活用も難しいという状況は、人文学の成果物のデジタル化の推進に対して負の影響を及ぼしている。

したがって、今後、論文発表や書籍出版における成果物のデジタル化や研究データ共有化の仕組みの構築を強く進めていく必要がある。これを通じて、人文学の研究成果の観光業などの他分野への応用や、人文学の研究成果を利用したシチズンサイエンスの活性化などが期待される。

また、人文学の研究成果が自然科学系の研究や社会課題解決に利活用される可能性もある。例えば、過去の地震の災害に関する研究データが今後の防災研究に生かされることや、あるいは過去の感染症に関する研究データが今後の感染予防に関する研究に生かされることがあり得る。最近の具体的事例としては、人間文化研究機構・国立歴史民俗博物館及び琉球大学の共同により新型コロナウイルスクラスターへの対応に関するアーカイブの試みと民俗学的調査が実施されている。これらのデータ蓄積が将来的な感染対策や感染者をどのように社会の中で受け入れていくかといった社会的なあり方を検討する材料になると期待される。また、研究資源のアーカイブ事業としては人社系データインフラストラクチャ構築

事業における東京大学史料編纂所の試みがあり、当該分野におけるデータ共有の促進を目的として実施されている<sup>31</sup>。本事業のようなデータ管理に関わるさまざまな取り組みを通じて、当該分野において、データ基盤を介した他分野でのデータ利活用や、情報技術によるデータ解析などの高度な研究データの利活用が促進されていくことが期待される。

このように人文学における研究データのデジタル化及びオープン化を進めていく上では、人文学分野の情報の性質に配慮する必要がある。例えば、肖像権やプライバシー権、差別を含む情報の取り扱いや、文化的・政治的な側面について配慮すべき点がある。この配慮すべき点は、時間的変化によって変わる点にも留意する必要がある。肖像権や歴史的な資料へと転化した情報等については、遺族等の関係者への配慮は必要であるものの、一般的に時間の経過とともにオープン化の範囲が広がる。その一方で、当時は問題ない表現として公開されていたものや、あるいは当時は問題なく共有されていた資料などが、その後の社会的倫理観の変化により、不適切な表現、あるいは共有が不適切な資料として扱われることもあるため、それらの情報を含む研究データの公開について留保が必要になる例もある。これらの情報について、その社会的情勢を適切に見極めながらオープン・アンド・クローズを実施する必要がある。

## (2) 農業分野におけるデータの管理・利活用

農業分野におけるデータマネジメントについては、内閣府・戦略的イノベーション創造プログラム（SIP）「次世代農林水産業創造技術」<sup>32</sup>において、農業に関わるさまざまなデータやシステムを連携させるための農業データ連携基盤（WAGRI）が開発され、現在は農業・食品産業技術総合研究機構（農研機構）が運営主体となり、本格稼働している。

WAGRI の基本的な利用方法は次のとおりである。まず、データ・システム提供者がWAGRIを通じて気象や農地、地図情報等のデータ・システムを提供する。次に、データ・システム利用者（農機メーカー、ICTベンダー等）はそれらのデータを使って、農業関連サービス（Webアプリケーション）を開発する。最終的に、農業者等はそれぞれの経営形態に合わせて戦略的にそれらのサービスを選択・活用し、データ駆動型農業を実施することが可能になる。

---

<sup>31</sup> 『人文系データインフラストラクチャ構築事業』、日本学術振興会 (<https://www.jsps.go.jp/j-di/torikumi.html>)

<sup>32</sup> 『次世代農林水産業創造技術』、内閣府 ([https://www8.cao.go.jp/cstp/gaiyo/sip/sympo1412/subject/subject\\_09.html](https://www8.cao.go.jp/cstp/gaiyo/sip/sympo1412/subject/subject_09.html))



具体的に WAGRI を介して提供されているデータやシステムは、地図、農地、気象、生育予測等に関わるものである。例えば、農研機構は土壌の種類や分布が分かるデジタル土壌図のデータを提供している。また、(株) ライフビジネスウエザーや (株) ビジョンテックはそれぞれ、農研機構の研究成果を基に、最長 26 日先までの日別気象情報 (1 km メッシュ) や水稲・小麦・大豆の生育予測システムを有料で提供している。

以上のように、WAGRI を介して、農研機構などのアカデミアや民間企業が生み出したデータやシステムが提供され、データ・システム利用者によるサービス開発を通じて、最終的には農業者等の便益へとつながる。この中で、上記の事例のように、データやシステムは有料で提供することも可能なため、データ・システム提供者のインセンティブにつながると期待される。また、データ・システム提供者は、自身が提供するデータの提供先や、提供項目を制限することができる。したがって、柔軟なオープン・アンド・クローズ戦略を実行可能な仕組みとなっている。

### (3) 自動運転における実験データの管理・利活用

SIP「自動運転 (システムとサービスの拡張)」(以下、SIP自動運転という)では、自動運転の実用化を高速道路から一般道へ拡張するとともに、自動運転技術を活用した物流・移動サービスを実用化することで交通事故低減、交通渋滞の削減、過疎地等での移動手段の確保や物流業界におけるドライバー不足等の社会的課題解決に貢献し、すべての国民が安全・安心に移動できる社会を目指している。そのための研究開発内容として、「自動運転システムの開発・検証 (実証実験)」、「自動運転実用化に向けた基盤技術開発」、「自動運転に対する社会的受容性の醸成」、「国際連携の強化」に取り組んでいる。この中で現在は、「自動運転システムの開発・検証 (実証実験)」で得られた実験データを「自動運転実用化に向けた基盤技術開発」へと提供し、利活用へとつなげる取組を実施している。

具体的には、「自動運転システムの開発・検証 (実証実験)」における実験データとしては、実験用の車両に搭載したセンサや信号機等のインフラ等から得られる測定データや、高精度3次元地図のデータが挙げられる。これらのうち、例えば車両に搭載したカメラで撮影した映像データ等が、「自動運転実用化に向けた基盤技術開発」における「自動運転の高度化に則したHMI 及び安全教育方法に関する調査研究」へと提供され、自動運転車と周囲交通参加者とのインタラクション・コミュニケーションの観測・分析が行われた。このような SIP自動運転施策の実施者間でのデータ提供及び利活用については、SIP自動運転の管理人であるNEDO (国立研究開発法人 新エネルギー・産業技術総合開発機構) との委託契約

のもとで、知財運営委員会によりマネジメントされ、実験データの提供、管理、利用等に係る規定に基づいて実施される。

またSIP自動運転においては、データの利活用により自動運転の実用化を促進するために、SIP自動運転施策の実施者間で議論するための共通の基盤とするとともに、SIP自動運転に参加していない企業や大学等の第三者へ提供し、フィードバックを得て、自動運転の実用化の加速に資することも検討している。その際には、データ管理者を誰にするべきかといったことや、どのようにデータを管理すべきか（データのカタログ化、文書化等）といったことが課題になると考えられる。また、データのフォーマットをいかに揃えるかといった点や、データ利用者が作成した派生データをいかに取り扱うかといった点も重要な課題になると考えられる。以上のような課題へ適切に対応していくため、SIP自動運転施策の実施者間で行うデータ提供、活用の事例を踏まえて、第三者へのデータ提供について検討しているところである。

## VI. 高度な研究データマネジメントを実現するための環境整備

### (1) 研究データマネジメント人材

研究データを適切に管理して利活用へとつなげていく主体は、その研究データを生み出した研究者自身に他ならない。したがって、研究者は研究データマネジメントを実践していくために求められる知識やスキルを身に付けていく必要がある。しかしながら、高度な研究データマネジメントの実現には、研究者だけではなく、研究者とともに協同で研究データマネジメントを実践する専門人材が必要不可欠である。特に、研究データを生み出す現場の研究者の負担を軽減することに加えて、研究データが「II 研究データに求められる要件」に記述された内容にそって管理がされていることを第三者的な立場から認証する仕組みや、そのための人材も重要になる。このような研究データマネジメント人材には、実験・観測データの体系的なマネジメントやデータキュレーション等の専門的な知見が必要である。

これまでに、いくつかの研究機関においては、研究活動に対して知識や経験を有するシニアの研究者や図書館業務に携わる人材を活用した実例がある。具体的には、NIMS においては、大型の国家プロジェクトでの DMP の策定・申請や、研究データマネジメント業務 (RDM) において、研究の全体像を理解できるシニアの研究者が指導的役割を担っている。また、上述の「IUGONET」エラー! ブックマークが定義されていません。プロジェクトにおいては、試行的ではあるものの、研究データへのメタデータの付与についてデータキュレーターとしての知識を持った図書館技術専門員との協働により、研究者の研究データマネジメントをサポート

ートした。さらに、シニアの研究者や図書館の人材以外の活用としては、大学院生や若手研究者に対して、教育・研究活動の一環として、研究データマネジメントの支援業務を担わせる可能性も考えられるが、この際には、単に従属的な活動にせず、大学院生や若手研究者のキャリア開発に有益なものとなるように配慮する必要がある。

後述するように、多くの大学で図書館のデジタル転換が行われつつあり、これらの活動と併せて研究データの管理や利活用を行うための専門人材の確保を行うことが必要である。特に、図書館は人文・社会科学の研究において伝統的に重要な役割を果たしており、新たな基本法の枠組みにおける人文・社会科学の推進にあたって、さらにデジタル転換や研究データマネジメント人材の育成を進めていく必要があると考えられる。

今後このような研究データマネジメント人材に関する取組みを広く定着させるためには、いくつかの重要な課題を解決する必要がある。例えば、まずそもそも研究データマネジメント人材に求められる要件を確立していく必要がある。この具体的内容については、研究分野にも依存するため、多様であると考えられるが、それも踏まえて検討していくべきである。その上で、研究データマネジメント人材育成の仕組みの確立や、雇用の安定化、異なる機関間の流動性も含めたキャリアパスの確立等が挙げられる。これらの人材は、教員とともに研究教育活動を担う専門人材として尊重され、適切な処遇が与えられることが必要である。また、研究データの管理及び利活用という業務自体を適切に評価する仕組みの確立が求められる。

## (2) 大学図書館による支援

上述のとおり、研究データの適切な管理や利活用においては、大学図書館の技術専門員と研究者による協同が試行的に実施された例があるが、さまざまな研究分野においてデジタル化が進み、データ駆動型の研究開発の重要性が高まる中、研究成果としての研究データの管理・利活用が重要となっていることから、大学図書館そのものの変革に対する期待は極めて大きい。すなわち、これまでは図書という知を集積して提供する「場」としての役割が中心であった図書館のあり方において、変容しつつある研究開発や教育のあり方に対応して、本当に必要な機能と不要な機能は何か、また新たに必要となる機能は何かについて見直すべきであると考えられる。またそれら機能を、場所に制限されることなく利用可能とするために、物理的な手段ではなくデジタル技術によってネットワークで処理するという視点も重要である。

新たに必要となる機能としては、例えば、研究データを管理するためのデータセンターとしての機能や、学部生や院生への研究データの利活用の仕方などを含めた研究のオリエンテーションやコンサルテーション機能、また、異分野を含めた共同研究の推進機能など、これまでの我が国の大学図書館ではあまり対応していなかった、より能動的な機能などが考えられる。このような大学図書館に寄せられる期待に対し、大学図書館を所管している文部科学省においては、大学図書館コミュニティと連携し、今後の大学図書館のあり方を検討する必要がある。

また、大学図書館への上記のような期待に応えていくためには、大学図書館人材の能力の向上も不可欠となると考えられる。例えば、欧米の大学図書館には専門職であるデータライブラリアンがおり、大学の学術情報知をいかに管理することが大学にとって適切であるかを議論している。すなわち、大学図書館における知の管理は高度な知識が必要な専門的な仕事であり、さらにこれからの図書館においては、デジタル技術にも一定程度精通した人材が必要となると考えられる。したがって、大学図書館員のそのような専門的な知識や技能の向上及び獲得のための環境整備や、新たな人材の配置が重要である。また、大学図書館員がそのような専門的な仕事を不自由なく実行できるように環境を整備していくことも求められ、国はそのための具体的な施策を検討し、積極的に支援する必要がある。

一方で、研究データは研究者における重要な知識資源であるとともに、大学経営層にとっては大事な経営戦略資源である。その意味においては、研究データの管理・利活用は、図書館だけが担うべき仕事ではない。したがって大学経営層は、研究データの管理・利活用における大学内の各部署が果たすべき役割を適切に検討し、実行させていくことが重要である。

以上のとおり、研究データの適切な管理・利活用を中心として、さまざまな分野における研究開発が変革していく中で、大学図書館への期待は非常に大きいところである。その期待に応えるためには図書館自身の変革も必要となってくる。政府としては、このような大学図書館の変革において、積極的な支援を実施していく必要がある。

## VII. 法的・制度的な課題

### (1) 研究データの帰属

研究データを管理・利活用していく上では、研究データの帰属に関する法的な定義が問題

になるが、「AI・データの利用に関する契約ガイドライン（2018年6月経済産業省）」<sup>33</sup>にあるように、いわゆるデータは無体物であり、日本の民法では所有権や占有権等の物権の対象とならない（民法206条、同法85条参照）。また、「データ・オーナーシップ」という言葉がよく使用されるが、これも同ガイドラインに説明されているとおり、現在のところ法的な定義がなく、一般的には、データに適法にアクセスし、その利用をコントロールできる事実上の地位や契約による債権上の地位を意味するものと考えられる。そして、データ創出への寄与度は事案によりさまざまであり、誰がデータ・オーナーシップを持つべきかという一律の基準を見出すことは困難であり、データ・オーナーシップがいずれの当事者に帰属するのかというオール・オア・ナッシングで交渉するよりも、個別事案に応じて「どのデータを、どちらの当事者が、どのような条件で利用できるのか」という利用条件をきめ細やかに調整し設定していくことが重要である。

以上のとおり、研究データを含めたデータ全般に関し、法的には原始的帰属は定められていない。したがって、利用条件を契約で決定していく必要がある。また、広く研究データを利用可能にするためにはクリエイティブ・コモンズ・ライセンス<sup>34</sup>のような運動が必要と考えられる。なお、法的な原始的帰属についての立法的な解決は、その必要性の検討を要する。

また、研究データの帰属の問題と関連して、特に公的資金により生み出された研究データの取り扱いについて、「委託研究開発におけるデータマネジメントに関する運用ガイドライン（平成29年12月経済産業省）」<sup>35</sup>では、研究データのうち公共性が高く、広範な利活用等を目的として委託者が管理する必要があるデータがあれば、公募の際に当該データを示した上で、委託契約において大学や企業など研究開発の委託を受けた受託者に対して、研究開発データの委託者への提供及び委託者におけるその自由な利活用の確保を約させるとしている。また、研究開発データの利活用を促進するため、受託者が自ら利活用又は第三者提供できることを原則とするとしている。このような研究開発データに関する取り扱いに関しても、法制度的な対応が必要かどうかは今後議論すべき問題であると考えられる。

---

<sup>33</sup> 『AI・データの利用に関する契約ガイドライン』、経済産業省（<https://www.meti.go.jp/press/2018/06/20180615001/20180615001.html>）

<sup>34</sup> 『creative commons』（[creativecommons.org](https://creativecommons.org)）

<sup>35</sup> 『委託研究開発におけるデータマネジメントに関する運用ガイドラインとナショナルデータカタログ』、経済産業省（[https://www.meti.go.jp/policy/innovation\\_policy/datamanagement.html](https://www.meti.go.jp/policy/innovation_policy/datamanagement.html)）

## (2) 研究データ授受に関する契約

研究データの授受に関する契約については、上述の「AI・データの利用に関する契約ガイドライン（2018年6月経済産業省）」<sup>33</sup>を参考にされたい。同ガイドラインの「データ編」においては、データに関する契約形態を、一方当事者から他方当事者へのデータの提供に関する「データ提供型」契約、複数当事者が関与して創出されるデータの取扱いに関する「データ創出型」契約、プラットフォームを利用したデータの共用に関する「データ共用型」契約の三種類の契約内容に分類し、それぞれにおける契約を締結する際の考慮要素、モデル契約等を説明している。

また、個別の業界における特殊性や課題を反映させた分野別の契約ガイドラインとして、「データの利用に関する契約ガイドライン産業保安版第二版（2019年4月経済産業省）」<sup>36</sup>と「農業分野におけるAI・データに関する契約ガイドライン（2020年3月農林水産省）」<sup>37</sup>が策定されている。これらについても必要に応じて参考にされたい。

以上のとおり、データ授受に関する契約ガイドラインについては、これまでに策定されているものがある。また今後も、特定の分野の特性に応じたガイドラインが策定されると期待される。また、研究データを広く利用可能とするにはクリエイティブ・コモンズ・ライセンス<sup>34</sup>のような活動も意義があると考えられる。

## (3) 研究データに個人情報を含む場合の取り扱い

我が国における個人情報保護法制は、個人情報の有用性に配慮しつつ、個人の権利利益を保護するものであるが、同じ研究者であっても所属機関（国立大学、公立大学、私立大学、国立研究開発法人、企業の研究所）によって適用される法令等が異なることもあり、研究者にとっては理解するのが難しいとの声がある。具体的には、我が国における個人情報保護法制において、民間部門（民間企業や私立大学が分類される）は「個人情報の保護に関する法律」（以下、「個人情報保護法」という）が、国の行政機関は「行政機関の保有する個人情報の保護に関する法律」が、独立行政法人等（国立大学が分類される）は「独立行政法人等の

---

<sup>36</sup> 『データの利用に関する契約ガイドライン産業保安版第二版』、経済産業省（[https://www.meti.go.jp/policy/safety\\_security/industrial\\_safety/sangyo/hipregas/files/20190425keiyakuGuideline.pdf](https://www.meti.go.jp/policy/safety_security/industrial_safety/sangyo/hipregas/files/20190425keiyakuGuideline.pdf)）

<sup>37</sup> 『農業分野におけるAI・データに関する契約ガイドライン～農業分野のノウハウの保護とデータ利活用促進のために～』、農林水産省（<https://www.maff.go.jp/j/kanbo/tizai/brand/keiyaku.html>）

保有する個人情報の保護に関する法律」が、地方公共団体・地方独立行政法人（公立大学が分類される）は「個人情報保護条例」がそれぞれ適用される。また、国立大学や国立研究開発法人については、EUの一般データ保護規則(GDPR)<sup>38</sup>の充分性認定から除外されている。これらにより、官民、地域を跨いだデータ利活用や、EUから日本への国境を越えたデータ移転の支障となる事例が各所で顕在化しつつあるとの指摘がなされてきた<sup>39</sup>。

さらに、海外における法令との関係に関しては、多国間で研究を行う際に、日本の法制度に加えて相手国の法制度を遵守することが求められるため、例えば、EU・GDPR、米国・カリフォルニア州消費者プライバシー法(CCPA)<sup>40</sup>等の海外の法制度を日本の研究者も理解する必要がある。

また、データの利活用を促進するため、2015年の個人情報保護法の改正に伴い匿名加工情報制度<sup>41</sup>が、2016年の行政機関の保有する個人情報の保護に関する法律及び独立行政法人等の保有する個人情報の保護に関する法律の改正に伴い非識別加工情報制度が、それぞれ導入された。匿名加工情報及び非識別加工情報は、特定の個人を識別することができないように個人情報を加工し、当該個人情報を復元できないようにした情報であり、通常の個人情報より緩やかな規律のもとで取り扱うことができる。アカデミアが保有する個人情報を含む研究データについて、産業界を含めて利活用を進めていくためには、アカデミアにおいて匿名加工情報又は非識別加工情報を作成して、産業界への提供を促していくことが、手法の一つとして考えられる。

匿名加工情報の作成等を公表している事業者の数は、制度創設以来着実に増えており、同制度の活用が進んでいる（2020年9月30日時点で579社<sup>42</sup>）。

---

<sup>38</sup> 『GDPR (General Data Protection Regulation : 一般データ保護規則)』、個人情報保護委員会 (<https://www.ppc.go.jp/enforcement/infoprovision/laws/GDPR/>)

<sup>39</sup> このような状況を踏まえ、我が国における個人情報保護法制の一元化について、内閣官房に設置された「個人情報保護制度の見直しに関する検討会」において検討が行われ、2021年通常国会に改正法案が提出された。

<sup>40</sup> 『California Consumer Privacy Act of 2018 米国「カリフォルニア州消費者プライバシー法 2018年」』、個人情報保護委員会 (<https://www.ppc.go.jp/enforcement/infoprovision/laws/CCPA/>)

<sup>41</sup> 『匿名加工情報制度について』、個人情報保護委員会 (<https://www.ppc.go.jp/personalinfo/tokumeikakouInfo/>)

<sup>42</sup> 『令和2年度上半期における個人情報保護委員会の活動実績について』、個人情報保護委

しかしその一方で、2020年度に総務省より報告された「情報通信白書令和2年版」<sup>43</sup>によると、企業側における匿名加工情報の提供や活用の状況について尋ねた設問に対し、我が国の企業の半数は「匿名加工情報について知らない」と答えており、今後、我が国において、産学での連携も含めて匿名加工情報の提供・活用を推進していく上では、このような匿名加工情報制度に対する認知度を一層向上させていく必要がある。

以上のような状況に鑑みると、研究者は、弁護士や個人情報の匿名加工の専門家などと、これら課題を解決するためのチームを作るなど、課題解決の仕組みを構築し、積極的にこれら課題に取り組むことが望ましい。例えば、大学の研究者がデータの加工を行う外部業者を容易に選定することができる環境を構築していくことも一案と考えられる。

また、研究データの利活用事例を積み上げて、当該事例を広く共有していくことも重要である。具体的には、例えば、ムーンショット型研究開発制度においても、研究開発を所管する政府関係部局は、各プログラムにおける研究者と積極的に対話をし、それぞれの課題と解決策について、広く共有し、他研究事業への横展開を促していくことが望ましい。

## VIII. 国際連携

### (1) 国際協調

近年、研究データの管理・利活用に関わる国際的な動きが活発化しており、欧州における EOSC<sup>14</sup>の構築など、さまざまな国や国際機関等において、研究データの管理・利活用を実現するための基盤やポリシー、ガバナンス等に関する取組が進められつつある。また、COVID-19の世界的な感染拡大は、WHOを中心とした研究データの国際的な共有化の潮流を作り出している。したがって、我が国としては、公的資金により生み出された国内の研究データの管理・利活用をさらに進めるためにも、国際的な取組と戦略的に連携し、我が国のプレゼンスの維持向上を実現することが重要である。

同時に、自由で開かれた研究活動を尊重し、我が国と価値観を共有する国や機関との協力関係の強化が重要である。例えば、研究データマネジメントの取組が進んでいる欧州や COVID-19のデータ共有の中心となっている米国と連携を取るだけでなく、アジア・太平洋地域でイニシアチブをとることで、我が国として存在感のある国際連携が構築できると考えられる。

---

員会 ([https://www.ppc.go.jp/files/pdf/R2\\_kamihanki.pdf](https://www.ppc.go.jp/files/pdf/R2_kamihanki.pdf))

<sup>43</sup> 『情報通信白書令和2年版』、総務省 (<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r02/pdf/index.html>)



このような国際的戦略を実現するためには、さまざまな国際会合等の場で、各国の戦略を把握しながら我が国の立ち位置を確立することが重要である。そのため我が国は、G7 茨城・つくば科学技術大臣会合において設置されたオープンサイエンス WG において、EU とともに共同議長を務めている。また、本年、ユネスコがオープンサイエンス諮問委員会を立ち上げたが、我が国からも代表を派遣し、この活動に参画している。OECD においても研究データの取り扱いに関する活動が実施されており、2021 年 1 月、公的資金の投入された研究データのアクセスを促進するための勧告が出された<sup>44</sup>。我が国としては、以上の活動の動向も注視しながらこれらに積極的に参画し、我が国のプレゼンスを高めることが重要である。

### (2) 各国における研究データ基盤との連携

上記の国際連携を介して、統合イノベーション戦略 2020<sup>1</sup>にも記載されているとおり、諸外国の研究データ基盤とも連携して巨大な「知の源泉」を構築し、あらゆる者が研究成果を幅広く活用できる世界を構築することが目指すべき姿である。

これまでに NII は、欧州原子核研究機構 (CERN) や米国 COS (Center for Open Science) 等とともに NII Research Data Cloud の開発を進めてきており、EOSC と連携している。この他にも NII は、韓国やマレーシア、ミャンマーといったアジア諸国やアフリカにおける研究データ基盤の構築・運用に関わる支援を実施することで、欧米以外の国々との連携も積極的に進めている。

これらの活動を通じて、各国のデータポリシーや技術、国際認証等の互換性を確保して、グローバルなデータプラットフォームが構築されることが期待される。

### (3) 各国間的な枠組み及び組織との連携

研究データマネジメントに関わる国際機関として以下のものが挙げられる。我が国としての研究データマネジメントを国際的に相互運用性のあるものにするためには、これら機関の動向を注視し、必要に応じて連携を取ることが重要である。

#### ・ Committee on Data for Science and Technology (CODATA)

国際科学会議 (ICSU) によって 1966 年に設立された学際的な科学委員会。科学と技術に関するあらゆるデータについて、その質、信頼性、管理、検索性の向上を行っている。

---

<sup>44</sup> 『Recommendation of the Council concerning Access to Research Data from Public Funding』、OECD Legal Instruments (<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>)

- Research Data Alliance (RDA)

2013年に欧州委員会 (EC)、米国国立科学財団 (NSF)、国立標準技術研究所 (NIST)、及びオーストラリアイノベーション省によって開始された研究コミュニティ組織。その使命は、データのオープンな共有を可能にする社会的及び技術的な橋を構築すること。

- World Data System (WDS)

科学データの保存・活用のための国際事業 (2008年設立)。科学界の国連と言われる国際科学会議が実施。分野横断的・分散協調的に、最新のIT技術を活用して、長期的な科学データの保存・利活用を国際的に確立することを目指す。

- Confederation of Open Access Repositories (COAR)

国際的なリポジトリコミュニティのネットワーク。人材育成や互恵的な活動支援だけでなく、リポジトリの高度化に必要なメタデータや語彙の共通化、次世代リポジトリの技術仕様の策定、次世代出版モデルの構築と実践など、多岐にわたる活動を展開している。

- The Future of Research Communication and e-Scholarship (FORCE11)

デジタル時代の現代的な学術コミュニケーションを目指し、研究者、学術情報流通関係者、政策担当者など、多様なステークホルダーによる任意団体 (2011年設立)。FAIRデータ原則を生み出した。

- Go FAIR

FAIRのデータ原則を実施し、データを発見可能、アクセス可能、相互運用可能、再利用可能にすることを目的とした、ボトムアップのステークホルダー主導の自主管理イニシアチブ。Implementation Networks (入力) を通じて協力している個人、組織、組織にオープンで包括的なエコシステムを提供している。

#### (4) 国際的共同研究におけるデータ連携

国際的共同研究の取組の中での研究データの共有や公開も多数行われている。例えば、ライフサイエンスの分野では、哺乳類を対象にゲノム情報から網羅的に遺伝子の機能を解析する、理化学研究所主導の国際共同研究プロジェクト「FANTOM (Functional ANnotation Of Mammalian genome) プロジェクト」の研究成果及び解析に用いたデータやその解析結果は、理化学研究所の FANTOM ウェブサイトや国際塩基配列データベースコレクション (International Nucleotide Sequence Database Collection)、バイオサイエンスデータベースセンター (NBDC) の生命科学系データベースアーカイブ (LSDBArchive)、公開リポジトリサービスを通じて、全世界に無制限で公開されている<sup>45</sup>。その結果、FANTOM プロジェ

---

<sup>45</sup> 『FANTOM』、国立研究開発法人理化学研究所 (<https://fantom.gsc.riken.jp/jp/>)

クトの成果は、京都大学山中伸弥教授らによる iPS 細胞の開発に関する論文等、さまざまな論文で引用され、世界中の研究に大きく貢献している。

また、COVID-19 に関する研究データやサービスを共有する取組として、COVID-19 データポータルがある。欧州では、COVID-19 に関連する研究データのポータルサイトを比較的早期から構築し、国際的な協調体制の構築を進めてきた。日本も欧州の呼びかけに賛同し、国立情報学研究所と国立遺伝学研究所が中心となって、国内の関係機関を連携しながら、COVID-19 データポータル JAPAN を構築した<sup>46</sup>。研究者が、国内外に散在する研究データや研究ツールなどの情報にいち早くアクセスできるよう、適切な分類とリンク先の選定、解説の付与を実施した情報発信サイトとして機能している。

## IX. まとめ

政府としては、本 WG のフェーズ 1 及びフェーズ 2 の検討の結果を踏まえ、次期科学技術・イノベーション基本計画や各年の統合イノベーション戦略に反映していくことが望まれる。また、これまでの各種ガイドライン等を包含する政府全体としての研究データに関する基本的な考え方を分かりやすく整理し、示していくことが重要である。また、これらを実行するため、府省横断的な連携体制を構築することが不可欠である。さらに、研究データの適切な管理・利活用を促進する観点から、研究者、研究事業やプログラム、機関等の取組状況について、適宜モニタリングとその改善に関するフィードバックを実施することも重要であると考えられる。

---

<sup>46</sup> 『COVID-19 データポータル JAPAN』、国立情報学研究所([https:// covid19dataportal.jp/](https://covid19dataportal.jp/))

(別添1) 本報告書で紹介した事例のまとめ

(以下の事例に加えて、その他の研究データマネジメントに関する事例も含めて、本報告書の図表集にまとめたので参照されたい)

事例名称等	内容	URL
IUGONET	分散管理されている太陽・地球物理に関する研究データの異分野間での流通を進め、新たな課題研究を推進することを目的として、メタデータのデータベースをネットワーク上で共有し、異分野の研究者でも容易に各研究データの概要を知ることができ、かつ、それら研究データへアクセスできるシステムを構築。	<a href="http://search.iugonet.org/list.jsp">http://search.iugonet.org/list.jsp</a>
エネルギー革新に向けたMI基盤の構築	機能性高分子材料のデータ駆動型科学による開発加速という目標において、まず、拠点となる大学と企業が1対1で連携し、互いのデータベースによるデータ連携の有用性を確認した後、大学を核としたコンソーシアムを形成し、各社の協調領域を見極めた上で、大学のデータベースやインフラの共用化、また、新たなデータ取得のための材料開発を実施することを提言。	<a href="http://www.cocn.jp/report/5bd594907a344229192c6132546a52143c31bde8.pdf">http://www.cocn.jp/report/5bd594907a344229192c6132546a52143c31bde8.pdf</a>
NIMS を中核としたオープンプラットフォーム (MOP : Materials Open Platform)	化学メーカーの保有する高分子材料について、NIMS がマテリアルズ・インフォマティクスなどを用いて評価する。その結果、得られた基礎的な研究データは各化学メーカーへと共有され、新たな材料開発等へと展開される (化学 MOP)。NIMS は鉄鋼メーカーとも、鉄鋼業界におけるオープンプラットフォーム (鉄鋼 MOP) の取組を実施している。	<a href="https://www.nims.go.jp/news/press/2017/06/201706190.html">https://www.nims.go.jp/news/press/2017/06/201706190.html</a>
JOINUS (Joint Open Innovation of drug repository)	アステラス製薬 (株)、田辺三菱製薬 (株)、第一三共 (株) が作成したドラッグ・リポジショニング化合物ライブラリーを、公募によって選ばれた国内研究機関 (大	<a href="https://open-innovation-joinus.jp/joinus/">https://open-innovation-joinus.jp/joinus/</a>

tioning)	学、公的研究機関、企業等)の研究者が、独自の評価系で評価する。得られた研究データは各社へとフィードバックされ、これにより、ライブラリー化合物の新たな適応症が発掘され、新薬開発につながると期待される。	
武田薬品工業(株)と東北大学東北メディカル・メガバンク(ToMMo)による共同研究	一般住民の全ゲノムリファレンスパネルの充実と日本人集団に特徴的な遺伝子型の発見、及び全ゲノム情報と脳MRI画像を含む健康情報・医療情報の統合的解析による新薬・治療法の研究開発を目的とした共同研究。武田薬品工業は、資金提供を行う代わりに、同共同研究を通じて取得される全ゲノム情報への1年間の優先的アクセス権を獲得。その後、武田薬品工業とToMMoは、我が国のゲノム情報の医療分野での活用を推進するため、全ゲノム解析とこの解析結果の利活用を目的としたコンソーシアムの実現を目指している。	<a href="https://www.tohoku.ac.jp/japanese/2020/04/press20200414-03-tommo.html">https://www.tohoku.ac.jp/japanese/2020/04/press20200414-03-tommo.html</a>
リジェネロン社が主導するバイオ医薬品会社のコンソーシアム	米国のバイオテクノロジー企業であるリジェネロン社が主導するバイオ医薬品会社のコンソーシアム。同コンソーシアムでは、参画企業が優先的に研究データにアクセスできる期間が設定されている。	<a href="https://www.ukbiobank.ac.uk/2019/03/new-data-available-exome-sequence-data-on-50000-participants/">https://www.ukbiobank.ac.uk/2019/03/new-data-available-exome-sequence-data-on-50000-participants/</a>
Clinical Trial Data Sharing (CTDS)	臨床試験データは高い透明性が求められるべきという考えと、臨床試験データは公共の財産であり、公衆衛生の向上をもたらす研究に積極的に活用すべきとの考えに基づいて、主として欧米の製薬企業により、臨床試験で得られた被験者レベルの詳細な個別被験者データ(Individual Participant/Patient Data: IPD)の公開・共有が行われるようになった。我が国でも、臨床試験データの共有(Clinical Trial Data Sharing:	<a href="http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/ctds.pdf">http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/ctds.pdf</a>

	CTDS) という言葉が広く知られるようになり、活発に議論が行われるようになってきた。	
AMED による「産学連携による次世代創薬 AI 開発」プロジェクト	産学が保有する創薬研究における多面的で膨大なデータを広く集約することで、産学で利用可能な化合物設計 AI を開発することを目的としている。同プロジェクトでは、federated learning 技術を活用することで秘匿のままの研究データ提供を可能にしており、開発された AI は、その開発に貢献した参加企業に提供することで企業内での研究開発に活用される予定。	<a href="https://www.amed.go.jp/program/list/11/02/001_02-04.html">https://www.amed.go.jp/program/list/11/02/001_02-04.html</a>
MELLODDY (Machine Learning Ledger Orchestration for Drug Discovery)	federated learning 技術を用いて、各製薬企業が自社の研究データを秘匿のまま提供して、それらのデータを基に、正確かつ効率的な創薬を実現するための AI 技術の開発を推進している。	<a href="https://www.melloddy.eu/">https://www.melloddy.eu/</a>
農業データ連携基盤 (WAGRI)	WAGRI を通じて提供された気象や農地、地図情報等のデータ・システムを基に、データ・システム利用者 (農機メーカー、ICT ベンダー等) が、農業関連サービス (Web アプリケーション) を開発。農業者等はそれらのサービスを選択・活用し、データ駆動型農業を実施することが可能になる。	<a href="https://www.naro.affrc.go.jp/project/research_activities/laboratory/rcait/121064.html">https://www.naro.affrc.go.jp/project/research_activities/laboratory/rcait/121064.html</a>
FANTOM (Functional Annotation Of Mammalian genome) プロジェクト	哺乳類を対象として網羅的に遺伝子の機能を解析することを目的とした、理化学研究所主導の国際共同研究プロジェクト。同プロジェクトでは、研究成果及び解析に用いたデータやその解析結果は、理化学研究所の FANTOM ウェブサイト等を通じて、全世界に無制限で公開されている。	<a href="https://fantom.gsc.riken.jp/jp/">https://fantom.gsc.riken.jp/jp/</a>
COVID-19 データポータル	COVID-19 に関する研究データやサービスを共有する欧州を中心とした国際的な	<a href="https://covid19dataportal.eu/">https:// covid19dataporta</a>

タル	取組。日本も欧州の呼びかけに賛同し、国立情報学研究所と国立遺伝学研究所が中心となって、COVID-19 データポータル JAPAN を構築。	l.jp/
----	--	-------

(別添2) WG 委員リスト

氏名 (敬称略・50音順)	所属・役職	備考
家 泰弘	独立行政法人 日本学術振興会 理事	
小賀坂 康志	国立研究開発法人 科学技術振興機構 未来創造研究開発推進部長	
落合 弘之	国立研究開発法人 新エネルギー・産業技術総合開発機構 技術戦略研究センター 標準化・知財ユニット長	
喜連川 優	大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 所長 国立大学法人 東京大学 生産技術研究所 教授	主査
後藤 文郷	住友化学株式会社 先端材料開発研究所 グループマネージャー	
末吉 互	KTS 法律事務所 弁護士	
谷 幹也	日本電気株式会社 研究・開発ユニット 上席技術主幹	
林 和弘	文部科学省 科学技術・学術政策研究所 科学技術予測センター 上席研究官	主査代理
松井 啓之	国立大学法人 京都大学 経営管理大学院 教授	
眞野 浩	一般社団法人 データ流通推進協議会 代表理事	
水野 充	国立研究開発法人 日本医療研究開発機構 基盤研究事業部長	
村山 泰啓	国立研究開発法人 情報通信研究機構 戦略的プログラムオフィス研究統括	
山地 一禎	大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 教授	