

回答

研究DXの推進—特にオープンサイエンス、
データ利活用推進の視点から—に関する審議
について



令和4年（2022年）12月23日

日 本 学 術 会 議

この回答は、内閣府からの審議依頼を受けて、日本学術会議に設置したオープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会、同オープンサイエンス企画分科会及び同オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会が中心となり審議を行ったものであり、日本学術会議として公表するものである。

日本学術会議オープンサイエンスを推進する
データ基盤とその利活用に関する検討委員会

委員長	喜連川 優	(連携会員)	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、東京大学特別教授
副委員長	林 和弘	(連携会員(特任))	文部科学省科学技術政策研究所データ解析政策研究室長
幹事	三枝 信子	(第三部会員)	国立研究開発法人国立環境研究所地球システム領域領域長
	西田 真也	(第一部会員)	京都大学大学院情報学研究科教授
	溝端佐登史	(第一部会員)	京都大学名誉教授・経済研究所特任教授
	小林 武彦	(第二部会員)	東京大学定量生命科学研究科教授
	小安 重夫	(第二部会員)	国立研究開発法人理化学研究所理事
	寺崎 浩子	(第二部会員)	名古屋大学未来社会創造機構特任教授
	仁科 弘重	(第二部会員)	愛媛大学学長
	藤原 康弘	(第二部会員)	独立行政法人医薬品医療機器総合機構理事長
	相澤 清晴	(第三部会員)	東京大学大学院情報理工学系研究科教授
	大橋 弘美	(第三部会員)	古河電気工業株式会社シニア・フェロー
	筑本 知子	(第三部会員)	中部大学超伝導・持続可能エネルギー研究センター教授
	菱田 公一	(第三部会員)	明治大学研究・知財戦略機構特任教授
	秋葉 澄伯	(連携会員)	弘前大学研究教授
	木部 暢子	(連携会員)	大学共同利用機関法人人間文化研究機構長
	澁澤 栄	(連携会員)	東京農工大学卓越リーダー養成機構特任教授
	高木 利久	(連携会員)	富山国際大学学長
	永井由佳里	(連携会員)	北陸先端科学技術大学院大学理事・副学長
	大向 一輝	(連携会員(特任))	東京大学大学院人文社会系研究科附属次世代人文学開発センター准教授
	穴戸 常寿	(連携会員(特任))	東京大学大学院法学政治学研究科教授

日本学術会議オープンサイエンスを推進する
データ基盤とその利活用に関する検討委員会
オープンサイエンス企画分科会

委員長	喜連川 優	(連携会員)	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、東京大学特別教授
副委員長	林 和弘	(連携会員(特任))	文部科学省科学技術政策研究所データ解析政策研究室長
	溝端佐登史	(第一部会員)	京都大学名誉教授・経済研究所特任教授
	武田 洋幸	(第二部会員)	東京大学執行役・副学長、大学院理学系研究科教授
	菱田 公一	(第三部会員)	明治大学研究・知財戦略機構特任教授

日本学術会議オープンサイエンスを推進する
データ基盤とその利活用に関する検討委員会
オープンサイエンス企画分科会
オープンサイエンス・データ利活用推進小委員会

委員長	喜連川 優	(連携会員)	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、東京大学特別教授
副委員長	林 和弘		文部科学省科学技術政策研究所データ解析政策研究室長
	溝端佐登史	(第一部会員)	京都大学名誉教授・経済研究所特任教授
	狩野 光伸	(第二部会員)	岡山大学副理事・学術研究院ヘルスシステム統合科学学域教授
	小安 重夫	(第二部会員)	国立研究開発法人理化学研究所理事
	武田 洋幸	(第二部会員)	東京大学執行役・副学長、大学院理学系研究科教授
	菱田 公一	(第三部会員)	明治大学研究・知財戦略機構特任教授
	木部 暢子	(連携会員)	大学共同利用機関法人人間文化研究機構長
	永井 良三	(連携会員)	自治医科大学学長
	生貝 直人		一橋大学大学院法学研究科教授
	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オープンサイエンス基盤研究センターセンター長・教授

本提言の作成に当たり、以下の方々に御協力いただいた。

田辺 新一	(第三部会員)	早稲田大学創造理工学部建築学科教授
大波 純一		大学共同利用機関法人情報・システム研究機構 国立情報学研究所オープンサイエンス基盤研究 センター特任准教授
大林 茂		東北大学流体科学研究所教授
小野 智弘		株式会社 KDDI 総合研究所 Human- Centered AI 研究所所長
金澤 輝一		大学共同利用機関法人情報・システム研究機構 国立情報学研究所コンテンツ科学研究系准教授
上東 貴志		神戸大学計算社会科学研究所センター経済経営 研究所教授
阪本 拓人		東京大学大学院総合文化研究科教授
桜田 一洋		慶應義塾大学医学部石井・石橋記念講座(拡張知 能医学)教授
高橋 恒一		特定国立研究開発法人理化学研究所生命機能科 学研究センター チームリーダー
谷藤 幹子		大学共同利用機関法人情報・システム研究機構 国立情報学研究所オープンサイエンス基盤研究 センター特任研究員
西岡 千文		大学共同利用機関法人情報・システム研究機構 国立情報学研究所コンテンツ科学研究系助教
一杉 太郎		東京大学理学系研究科化学専攻教授
藤嶋 誠		DMG 森精機株式会社取締役副社長
船守 美穂		大学共同利用機関法人情報・システム研究機構 国立情報学研究所情報社会相関研究系准教授
南山 泰之		大学共同利用機関法人情報・システム研究機構 国立情報学研究所オープンサイエンス基盤研究 センター特任助教
三輪 泰史		株式会社日本総合研究所創発戦略センターエク スパート

本回答の作成に当たり、以下の職員が事務及び調査を担当した。

事務	佐々木 亨	参事官(審議第二担当)
	高橋 直也	参事官(審議第二担当)付参事官補佐
	齊藤 美穂	参事官(審議第二担当)付審議専門職
	松本 拓馬	参事官(審議第二担当)付審議専門職付
調査	奥和田久美	上席学術調査員

要 旨

令和4年3月23日付け府総第104号-2、府科事第344号-2の審議依頼（研究DXの推進—特にオープンサイエンス、データ利活用推進の視点から—に関する審議について（依頼））を受けて、日本学術会議では、オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会オープンサイエンス企画分科会にオープンサイエンス・データ利活用推進小委員会（以下、「小委員会」という。）を設けて、議論を行った。

日本学術会議ではオープンサイエンスに関連して、2016年と2020年に提言を発出し、オープンサイエンスの推進に当たっての課題を整理している。それらの提言以降も、オープンサイエンス及びデータ駆動型科学は、多様な分野において進展し、より開かれた知識を柔軟に活用した新しい研究スタイルの確立や、社会課題解決に向けた連携が進みつつある。その一方で、この分野における我が国の取組は必ずしも十分とはいえない。データ駆動型科学を中心とした新しい研究基盤の構築に向けて、欧米の主要国から遅れをとることに、学术界としても危機感を抱いている。

小委員会ではこのような認識から、多くの分野におけるデータ駆動型科学の先進的実践に関し事例を聴取し、我が国の現況を把握するとともに、データ駆動化を加速するための要件について議論し、聴取したような事例が他の分野においても広く実践・応用されるためには、どのような方策と体制整備が必要であるかという観点から、オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会、同委員会オープンサイエンス企画分科会及び小委員会において審議を重ねた。

(1) 研究データの共有・公開も含めたオープンサイエンスに対する日本学術会議としての考え方

日本学術会議は、第四の科学と呼ばれるデータ駆動型科学が近年広く進展し多様な成果を生み出す中で、その中核ともいえる研究データの共有・公開を推し進めようとするオープンサイエンスの考えに強く賛同するものである。第三の科学である計算科学は方程式によって記述可能な世界を対象とし、スーパーコンピュータ（以下「スパコン」という。）を駆使することにより科学に大きな進展を与えたものの、地球、身体、言語のような方程式を明らかにすること自体が困難と考えられる複雑な対象も多々存在する。これらの学理の解明に挑戦すべく、データ駆動型科学が提唱され、今日、多くの分野で推進されるに至っている。「研究DX」なる用語は、研究にデジタル技術を積極的に活用することにより科学の飛躍的進展を目指そうとする多様な取組の総称といえる。その中でも、研究データを科学者コミュニティで積極的に共有し、最大限利活用することを目指すオープンサイエンスは最も重要な取組といえる。データ駆動型科学は学術全体において大きな潮流となりつつあり、小委員会では、先端的なデータ利用を推進する研究者等から集中的に現況と課題の聴取を進めた。例えば、物質材料学ではロボットを導入し、実験の自動化によりデータの取得を著しく効率化するとともに、得られたデータにAIを適用して新たな仮説を生み、途切れることなく実験を可能とすることで新素材の開発速度を飛躍的に加速するなど、科学実験のスタイルが大きく変容しつつある。この手法はAutomated Research Workflow（以下「ARW」という。）と呼ばれている。データは科学的発展だけではなく人類の社会生活にも多大な恩恵を与えている。例えば、ウェブ上の膨大な文書を利活用することにより1億超の対

訳テキストが整備され、これを基に高度な言語処理技術を駆使して構築された高品質翻訳サービスが世界の多言語コミュニケーションに革命をもたらしている。そもそも、このような言語処理技術の進展は、言語学を飛躍的に発展させる可能性を持ち、データは、人文・社会科学をさらに大きく進展させるものとして大きな注目を浴びている。このようにデータは多様な分野で大きな変革を創出しつつあるが、研究データの利活用には、データの取得、蓄積、検索、メタデータ付与、それらを統合するデータプラットフォームの構築、人材育成等多様な側面があり、各々に内在する課題を丁寧に明らかにし、解決していくことが肝要である。そのためには、大学等と研究機関に加えて、研究助成団体、学会、出版社等のステークホルダーとの調整も不可欠である。法的取扱いが不十分な無形資産としての研究データの扱いに関しては、産学連携においても、新たなルールが必要である。

以上のような日本学術会議としての問題意識の下、審議依頼の各事項に対して、データ駆動型科学やオープンサイエンスを推進すべく、以下の提案を含めて回答する。

(2) 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）

【提案1】研究者が容易に利用可能な研究データプラットフォームの構築

統合イノベーション戦略推進会議が2021年4月に公開した「公的資金による研究データの管理・利活用に関する基本的な考え方」では、研究データの取扱いに関して基本的な考え方が丁寧にまとめられている。「6 研究者の責務」には多くの指摘事項があるが、研究者がそれらの指摘に対応するための負担からデータ駆動型科学への参画を躊躇することとなっては本末転倒である。すなわち、データを扱う際の種々の手続に係る負担が限りなく小さくなるよう工夫をすることが肝要である。例えば、EUでは約10億ユーロの予算を投下し、デジタル技術を駆使した広範な研究開発により European Open Science Cloud の構築を進めている。統合イノベーション戦略推進会議は、国立情報学研究所（以下「NII」という。）の NII Research Data Cloud を中核的なプラットフォームとして位置付け、NIIはその開発を2017年より進めているが、より大規模な開発により加速することが不可欠である。

【提案2】データプロフェッショナルの育成と多面的な研究評価の実現

データ駆動型科学があらゆる分野において推進されるためには、データを的確に扱うための人材や関連のスキルが、研究の現場に浸透することが要となる。データを扱う多様なスキルを有するデータプロフェッショナルを機関戦略として配備する海外事例も見られるが、プールされたデータプロフェッショナルが全国の大学に協力できる仕組みも構築されることが望ましい。その際、複数の機関やメンバーからなるチームサイエンス¹が可能となるように、大学等には研究データガバナンスが確立している必要がある。また、「研究評価のためのパリ宣言²」を端緒に、論文業績以外の多様な研究貢献を評価し、チームサイエンスを実現することで、オープンサイエンスや産学官の協働を通じたイノベーションの創出、市民参加も含めた社会的課題解決につなげよ

¹ 異分野や異なるスキルを有する研究者等の連携による「チームサイエンス」というコンセプトの重要性は、全米アカデミーズの報告書[1][2]や研究評価改革のためのパリ宣言[3]等に言及されている。

² 本文における参考調査報告1-5を参照のこと。

うという動きも見られる。我が国においても同様の取組を通じて、学術の新たな展開につなげることが望まれる。

【提案3】 モニタリング機構に基づくデータ駆動型研究の不断の改善

研究データの他者との共有がどの程度進められているか、それによって成果が拡大しているか、データに関連するツールはどの程度利用されているか、データはどの程度容易に検索できているかなど、データ駆動型研究環境を可観測化し、不具合を常に監視するモニタリング機能の開発や運用を担う機構の整備が肝要である。そのためには、研究成果を体系的かつ長期的に、また分野横断的に学術知識情報として集約する体制が不可欠である。

(3) 各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例など）、データ共有への具体的取組方策（データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など）

【提案4】 研究自動化 (ARW) に向けた情報技術、計算資源の集約

ARW では、人工知能を始めとする先端的デジタル技術やロボット技術等を集約することにより科学的知見を獲得するプロセスの圧倒的加速化に関して手応えが得られつつある。第四の科学と呼ばれるデータ駆動型科学が大きく注目されているとはいえ、実験・観測科学、理論科学、計算科学といったこれまでの科学が色褪せるものではない。観測データに比してスパコンが生み出すデータ量は巨大であり、スパコン等の大型実験装置（コアファシリティ）を連成した大きなクローズドループの形成と、ループを成す ARW ごとのオープンデータ化により、研究全体の効率化を図ることができる。ARW での実験手順はソフトウェアとして研究者からアップロードされ、ロボット化された ARW 環境は研究者間で共有されることが期待される。このような先進的 ARW への戦略的研究開発支援も強く望まれる。

【提案5】 分野を越えた連携を実現する FAIR 原則³の追求

多様な分野間の連携においては論文だけではなくデータを介することにより ARW の更なる加速が期待されるが、いかなるメタデータ付与がデータの利活用に資するかは分野依存性が高い。FAIR 原則に基づき、各分野における積極的な分野間連携の実践が望まれ、また、その経験値の共有を可能とする取組が推奨される。日本学術会議や学会等において、FAIR 原則を追求する踏み込んだ議論が期待される。

【提案6】 法制度面でのデータガバナンスの構築

データ駆動型科学を推進していく上では、各種データの保護に関する法整備と、学術界としての法遵守や自主的な施策を通じて社会からの信頼を得ることが不可欠である。特に個人情報に関しては、改正された個人情報の保護に関する法律について学術機関として周知・遵守することはもとより、国内外の関係者から信頼を得られる個人情報保護の施策を進めながら、国際的なデータ共有に基づく共同研究を加速させていく必要がある。さらに、公的機関や企業の保有するデータや、健康・医療データの学術研究での活用円滑化を含め、社会全体でのデータ活用を推進するための新たな法的アプローチを検討していく必要がある。

³ 本文における 3 (3) ②を参照のこと。

目 次

1	はじめに	1
(1)	審議依頼	1
(2)	背景－巨大オープンデータと研究の自動化が突き動かす研究 DX の動き	1
2	国内外の動向と課題	2
(1)	国内の状況	2
(2)	海外の状況	5
(3)	日本の強みとこれから取り組むべき課題	7
3	審議依頼への回答	8
(1)	研究データの共有・公開も含めたオープンサイエンスに対する日本 学術会議としての考え方	8
(2)	大学・国立研究開発法人等において必要となる研究データ管理・利 活用のための課題の整理と具体的方策（管理・活用体制の整備方策、 人材確保・育成方策など）	9
①	【提案 1】研究者が容易に利用可能な研究データプラットフォームの 構築	9
②	【提案 2】データプロフェッショナルの育成と多面的な研究評価の実 現	11
③	【提案 3】モニタリング機構に基づくデータ駆動型研究の不断の改善	13
(3)	各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために 考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例な ど）、データ共有への具体的取組方策（データ共有へのインセンティブ 付与のための方策、分野間連携のためのコミュニケーションの在り方 など）	15
①	【提案 4】研究自動化 (ARW) に向けた情報技術、計算資源の集約	15
②	【提案 5】分野を越えた連携を実現する FAIR 原則の追求	17
③	【提案 6】法制度面でのデータガバナンスの構築	19
4	おわりに	20
<	参考調査報告 1 > 欧州の状況	22
<	参考調査報告 2 > 米国の状況	24
<	参考調査報告 3 > 豪州の状況	25
<	参考文献 >	29
<	参考資料 1 > 審議経過	38
<	参考資料 2 > 審議依頼公文	41
<	付録 1～9 >	43

1 はじめに

(1) 審議依頼

2022年3月に内閣府から次の審議依頼を受け、これまでの日本学術会議における検討及び各分野の多様性を踏まえ本回答をするものである。

- ① 研究データの共有・公開も含めたオープンサイエンスに対する日本学術会議としての考え方の取りまとめ
- ② 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）
- ③ 今後のデータ駆動型科学の振興のために考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例など）、データ共有への具体的取組方策（データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など）

(2) 背景－巨大オープンデータと研究の自動化が突き動かす研究DXの動き

デジタルトランスフォーメーション（DX）が各所で推進される中、科学においてもいかにデジタル技術を巧みに活用するかが肝となる。科学のデータ駆動化の進展により実験のスタイルが大幅に変容しつつある。東京大学の一杉太郎教授[4]（付録1参照）、国立研究開発法人物質・材料研究機構の松田翔一チームリーダー[5]（付録4参照）の研究室では、ロボットが休みなく繰り返す実験から生み出されるデータを活用し、新素材の発見を始めとする材料科学の進展を飛躍的に加速する成果が報告されている。現状のAIにとって「データ」は燃料として必須である（Data Fuel AI）。測定したデータを利用してAIが判読し、次の実験計画を立案する自律的サイクルが繰り返される方法論が、材料科学や生命科学領域で生まれつつある[6]。AIが使うデータを人手で生成するのに比べ、ロボットを駆使することにより、100倍以上（桁違い）の高速化が見込める。生み出されたデータを適切に管理し、他者でも利用可能とする、すなわちデータをオープンにし、多くの研究者間で共有することにより更に発見が加速することは自明であろう。先端的なデジタル技術の利活用を可能とする大規模データプラットフォームは不可欠であり、その構築は急務といえる。

従来から高解像度・高性能な電子顕微鏡等の高額な実験装置（以下「コアファシリティ」という。）の共用は進められていたが、皮肉にもコロナ禍により、測定データのサイバー空間での共用、実験の遠隔化が加速されることとなった。物理的な実験だけではなく、昨年のもろとみ博士のノーベル賞受賞に見られるスーパーコンピュータ（以下「スパコン」という。）の世界も同様である。

人類にとって今日最大の課題の一つともいえる気候変動予測に対し、文部科学省の気候変動予測先端研究プログラムや環境省の気候変動影響予測・適応評価の総合的研究が推進されているが、それらの研究によって生み出され

るデータは約 10～15 ペタバイトに達すると見込まれている。研究者はスパコンが生み出す超巨大なデータをもはや自分のコンピュータにダウンロード不能であることから、自在に多様な解析を可能とするデータプラットフォームは必須であり、気候変動の分野に留まらず多くの研究分野で極めて重要な役割を果たす。

上述のような取組は自然科学や理工分野における巨大なサイエンスだけではなく、人文・社会科学分野においても大きなインパクトをもたらし、巨大データを利用した研究が生まれつつある。ウプサラ大学（スウェーデン）における紛争データプログラムは、長年にわたり武力紛争に関する情報を収集し、多数の研究者のデータプラットフォームの役割を担っている。我が国では、東京大学の阪本拓人教授らが国際連合の安全保障理事会における 30 年にわたる公開議事録データを、コンピュータを使って分析することにより、主要国の脅威認識の変化や異同の把握を試みている（付録 9 参照）。

このような状況下において、日本学術会議ではオープンサイエンスに関連して、2016 年[7]と 2020 年[8]に提言を発出し、オープンサイエンスの推進に当たっての課題を整理している。オープンサイエンスは、オープンアクセスに始まる様々な研究プロセスの公開を含んでおり、その中でもオープンデータに注目が集まってきた。先の提言以降も、上記で述べたように、オープンサイエンス及びデータ駆動型科学は、多様な分野において浸透し、より開かれた知識を柔軟に活用し、あるいは、研究者間の共創による新しい研究スタイルの確立が進んでいる。また、社会課題解決に向けて、市民を含む多様なステークホルダーによる連携も進んでいる。その一方で、これらの分野における我が国の取組は必ずしも十分とはいえない。データ駆動型科学を中心とした、新しい研究基盤の構築に向けて、欧米の主要国から遅れをとることに、学术界としても危機感を抱いている。

このような認識から、オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会オープンサイエンス企画分科会の下に設置したオープンサイエンス・データ利活用推進小委員会（以下「小委員会」という。）では、多くの分野におけるデータ駆動型科学の先進的実践に関して事例聴取し、我が国における先端的データ駆動型研究の現況を把握した。その上で、データ駆動化を加速するための要件について議論するとともに、聴取したような事例が広く実践・応用されるためには、どのような方策と体制整備が必要であるかという観点から、オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会、同委員会オープンサイエンス企画分科会及び小委員会において審議を重ねた。

2 国内外の動向と課題

(1) 国内の状況

2016 年に日本学術会議オープンサイエンスの取組に関する検討委員会が公

開した提言「オープンイノベーションに資するオープンサイエンスのあり方に関する提言」では、国レベルで研究データ基盤を整備する必要性を指摘した[7]。国立情報学研究所（以下「NII」という。）では、文部科学省からの支援を受け、オープンサイエンスを推進するための汎用的な研究データ基盤 NII Research Data Cloud（以下「NII RDC」という。）の開発を2017年から開始し、2021年からサービスの本格運用を実施している。同サービスは、統合イノベーション戦略推進会議によって、我が国におけるオープンサイエンスを推進するための中核的プラットフォームとして位置付けられている。

NII RDCは、「管理基盤」、「公開基盤」、「検索基盤」の3つの基盤から構成される。公開基盤と検索基盤は、オープンアクセスを支援するインフラとしてこれまでNIIが提供してきた、機関リポジトリ⁴のクラウドサービス JAIRO Cloudと文献検索サービス CiNii を、論文や書籍、博士論文だけでなく、研究データも扱えるように拡張することで実現している。公開基盤である JAIRO Cloudは、主として大学や研究機関の図書館が運用する機関リポジトリとして活用されており、2022年11月の時点で既に国内の748機関が利用している。検索基盤である CiNii Researchは、年間6億ページビューを超える従来の文献検索サービス CiNii Articlesに加え、書籍を対象とする CiNii Books、博士論文を対象とする CiNii Dissertation、科学研究費補助金の研究プロジェクトを対象とする KAKEN の情報を全て統合すると同時に、研究データを対象とする国内のデータベースの情報も集約する巨大な学術知識情報を検索の対象とする。2022年11月の時点での収録総コンテンツ数は、1.8億件を超えている。データベースの収録数は、順次増加している。

公開基盤と検索基盤が公開された研究成果を対象とするのに対し、管理基盤である GakuNin RDMは、研究中の非公開状態での研究データの管理と共有を支援する新しいサービスである。公開基盤における共有と、日々の研究の過程で実現したい共有では、必要となる機能が異なる。NII RDCは、研究者のニーズを満足させるために3つのサービスを分離しつつも、それらの基盤を有機的につなげることにより、オープンサイエンス時代の研究ワークフローを支える環境の構築を目指している。GakuNin RDMは、2021年度から運用を開始したばかりであるが、2022年11月の段階で51機関が利用を開始している。各機関では、まだまだ限定的な利用ではあるが、機関として研究者に提供する研究データの管理や共有のための共通基盤として順調に成長している。

2020年に日本学術会議オープンサイエンスの深化と推進に関する検討委員会が発出した提言「オープンサイエンスの深化と推進に向けて」では、NII RDCの更なる高度化が必要であることが指摘され[8]、それを受けた研究開発が現在進められている。2022年からは、文部科学省による「AI等の活用を推進す

⁴ 大学や研究機関がその所属研究者の知的生産物を電子的形態で集積し保存・公開するためのウェブサービス。

る研究データエコシステム構築事業」の支援を受けて、図1に示すNII RDCの利活用を促進するための機能の高度化と、全国の大学や研究機関で基盤の利用を促進するための事業が展開されている[9]。この事業では、単にNII RDCに必要とされる機能の研究開発をするだけではなく、表1に挙げるような取組も同時に行うことにより、オープンサイエンスやデータ駆動型研究の推進に寄与するための必要な環境を整備することを目的としている。この中でも「ルール・ガイドライン整備」においては、個人情報等の研究データを取り扱う上での留意等についても対象とすることが事業の実施範囲として定められている[10]。こうした取組のほかにも、NIIでは、日本学術会議との連携の下、「オープンサイエンスのためのデータ管理基盤ハンドブック」を作成・公開しており[11]、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上での素地は、着々と整ってきている。

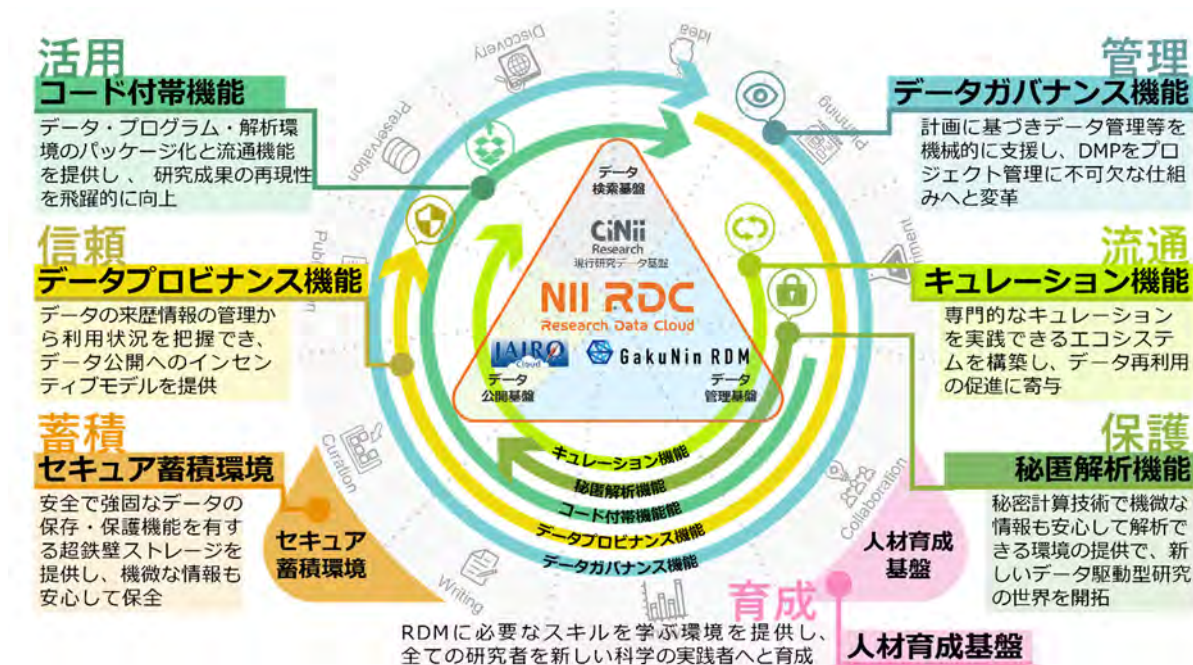


図1 NII Research Data Cloudが新たに実装する高度化機能群
(出典) NII「AI等の活用を推進する研究データエコシステム構築事業[9]」

表1 AI等の活用を推進する研究データエコシステム構築事業の実施内容

チーム名	実施内容
プラットフォーム連携	各研究分野や機関が運用するリポジトリやデータプラットフォームを相互接続するためのルール作りや規定整備
融合・利用開拓	異なる分野間でのデータ連携を実現するためのシーズやユースケースの創出
ルール・ガイドライン整備	研究データの効果的かつ効率的な利活用のためのルールやガイドラインの策定や標準化

人材育成	研究データ管理に必要な人材育成に必要となるスキルセット等の整理や教材開発
------	--------------------------------------

(出典) NII「AI等の活用を推進する研究データエコシステム構築事業[9]」

(2) 海外の状況

オープンサイエンスを推進するための環境整備という観点では、欧州の取組が象徴的である。欧州では、デジタル単一市場戦略に基づきオープンサイエンス政策を2016年から精力的に推し進め、その実現に向けた体制強化や制度設計が進んでいる。汎欧州型のオープンサイエンス基盤（European Open Science Cloud: EOSC。以下「EOSC」という。）[12]の構築が政策の中心にあり、欧州諸国の既存の基盤との連携を通じた、分野横断的な研究データの共有や再利用に取り組んでいる。付随する標準化、人材育成、法制度への準拠に関わるサービス群は、汎欧州のデータ駆動型科学に向けた「統合環境」の形成を担っている。2021～2027年のプログラムの予算総額は、約10億ユーロと発表されており[13]、先出の文部科学省「AI等の活用を推進する研究データエコシステム構築事業」における5年間の予算総額の約28倍の支援規模である。あわせて、EOSCとほぼ同規模の予算が、欧州研究基盤戦略フォーラム（European Strategy Forum for Research Infrastructure: ESFRI。以下「ESFRI」という。）[14]を通じて、主要な研究設備等に投下されている。そこで構築される研究設備やサービスもEOSCに有機的に統合される予定であり、分散協調型の研究基盤の構築が着々と進んでいる。

欧州では、さらに学術システムを根底から見直す研究評価改革が進行している。近年の学術は、オープンサイエンスにおける第三者のデータ利用や、データ駆動型科学を強力に推進するために、個人研究からチームサイエンスへと移行している。チームサイエンスを成り立たせるためには、最終的な研究成果となる学術論文だけでなく、多様なチームメンバーの貢献を評価する必要がある。欧州では、2022年2月の「研究評価のためのパリ宣言」[3]において「研究のオープン化が、研究の質、効率性、インパクトにつながり、チームサイエンスを醸成する」ことを確認した上で、研究評価改革の方向性を示し、40か国350機関の関心表明の下、2022年7月に「研究評価の改革に関する合意」[15]を取りまとめ、同年12月には、欧州委員会、欧州大学協会（以下「EUA」という。）、欧州研究助成機関等の団体であるScience Europeが中心となって研究評価促進連合（以下「CoARA」という。）[16]を正式に設立した。EOSCの成果として提供されているオープンサイエンスのモニタリング基盤を通じて、研究者の多様な貢献の評価が部分的に開始されている。研究の質やインパクトの向上につながる研究評価改革は、「欧州研究圏政策アジェンダ2022-2024」にも位置付けられ[17]、欧州各国レベルの政策の方向付けにも大きな影響を及ぼしており、汎欧州の大きな潮流へとつながっている（参考調査報告1-5参照）。

欧州においてオープンサイエンス政策が精力的に推進される一方、米国においてはこれに関係する動きがあまり見られなかった。米国国立科学財団（以下「NSF」という。）において先進的サイバーインフラ室⁵が設置され、関連の助成がなされたり（参考調査報告2-3参照）、オープンサイエンスや研究の再現性等について米国科学アカデミーから提言等はなされたりしているものの、連邦政府からの動きがなく、国全体としてのムーブメントにつながっていなかった。

そのような中、2022年8月末、米・科学技術政策局（以下「OSTP」という。）が公的資金を得て生み出された研究成果について「即座オープンアクセス」（以下「即座OA」という。）の方針を打ち出し、世界中の注目を集めた[18][19]。これが、データ駆動型科学やオープンサイエンスの必要とする研究評価改革につながるかは不明であるが、オープンに流通する研究データが飛躍的に拡大することは間違いなく、そこからデータ駆動型科学が自然と誘発されることが推測される。ただし、米国の「即座OA」方針の施行は2025年以降になると予想され、欧州に比べると歩みが遅いといえる⁶。

豪州は、オープンサイエンスという観点では大きな動きは見えないが、研究データガバナンス及びデータ駆動型科学を推進するための体制整備に関して、機関による積極的な取組が多数見られる。その原動力として、豪州研究データコモンズ（Australian Research Data Commons: ARDC。以下「ARDC」という。）[20]が、研究データに関わる豪州の研究競争力強化について旗振り役となっていることが挙げられる。データリポジトリを提供するだけでなく、研究データに係る課題やガイドラインの整備、人的ネットワークの形成、スキル開発等も担い、関連する学术界や産業界を統括するコミュニティの中心となっている（参考調査報告3-2参照）。ARDCだけでなく、豪州研究助成機関が取りまとめた「責任ある研究実践のための行動規範2018」や、豪州の研究設備に関わる「豪州協働型研究基盤戦略（National Collaborative Research Infrastructure Strategy: NCRIS。以下「NCRIS」という。）」[21]、大学のコアファシリティ等に対する機関横断的な支援体制の整備等、豪州の多様な領域の政策が整合性をもっている点も注目に値する（参考調査報告3-3参照）。

豪州では、個別の大学において、研究担当理事を中心とした体制の強化が進んでいる。深刻な情報流出の問題に起因し、現在では、研究データを大学の提供するストレージにしか保存できないルールを持つ機関もある（参考調査報告3-1参照）。データ駆動型科学の推進という観点からは、機関戦略としてリサーチエンジニアをプールし、学内の個別の研究プロジェクトに対して、大型計算機や学内研究設備の利用に関する指導や助言を行うといった丁寧な支援

⁵ 米国国立科学財団によって設立された、最先端の研究基盤、研究ツール等の開発や提供を支援するための組織。

⁶ 欧州では「即座OA」を求めるプランSが2018年9月発表され、2021年1月に発効している（参考調査報告2-1参照）。

体制が整ってきている。以上のような国及び各機関における戦略性が、研究力強化に向けての豪州全体の大きな力につながっている（参考調査報告 3-5 参照）。

(3) 日本の強みとこれから取り組むべき課題

EOSC では、従来から構築と運用が進められてきた研究基盤を有機的に連携させることを主眼としているが、小規模なサービスには支援が届いていない。例えば、大学が運用する機関リポジトリは、個別の機関の自助努力で管理されており、そのことがシステムの持続的な運用や高機能化への大きな妨げになっている。それに対して、我が国では、JAIRO Cloud による一括したシステムの運用や機能強化が実施されており、海外の状況と比較しても大きな優位性を保っている。研究中のデータ管理を受け持つ GakuNin RDM についても、2021 年の運用開始から順調に利用機関数が増加している。総合大学だけではなく、最近では理工系、人文社会科学系、医学系の単科大学からの利用が増加しているという特徴がある。こうした共通基盤を NII RDC として提供し、国全体として効果的に基盤の研究開発や運用が実施できている点は、世界からも高く評価されている。特にこれからリポジトリの普及を目指している発展途上国からは、日本のモデルを採用しようとする事例も出てきている[22][23]。共通基盤の整備が進む中、オープンサイエンスやデータ駆動型研究の進展につなげるためには、これを全国の大学や研究機関に普及させる必要がある。共通基盤の整備自体は、日本の優位性が保てている一方で、海外との状況を総合的に比較した場合には、日本として取り組むべき以下のような課題が浮き彫りになってくる。

NII RDC は、汎用的な研究データ基盤として大学や研究機関での採用が進んでいるが、研究分野における本格的な利用はこれからである。文部科学省による新事業により、分野プラットフォームとの連携やその事例の創出が試みられているが、EOSC との比較からは、国全体に広がる更なる大規模な試みが必要である（提案 1）。この基盤を活用した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ためには、人材の育成が不可欠であることは言うまでもない。海外では、データプロフェッショナルと呼ばれる人材を強化し、チームサイエンスという活躍の場を創出しながら研究力の強化に努めている。インパクトファクターのみに頼らない研究評価の抜本的な見直しの下、研究者にデータ共有や公開のインセンティブを与える仕組み作りにも取り組んでいる。そうした新たな研究活動を見える化するためのモニタリング機能の開発や運用を担う機構の整備についても積極的な検討が進められている。これらは我が国として今後取り組んでいくべき課題であり（提案 2 及び提案 3）、審議依頼 2 に対する日本学術会議としての提案内容として、次章以降でまとめる。

小委員会では、先端的なデータ利用を推進する研究者等から集中的に事例の聴取を進めてきた。その中でもいくつかの事例で見られた、研究自動化

(Automated Research Workflows: ARW。以下「ARW」という。)と呼ばれる新たな研究手法により、科学実験のスタイルは大きく変容しつつあることが鮮明になってきた。研究データ基盤が、この新しい研究のスタイルを積極的に支援することは、これからのデータ駆動型科学の振興のために不可欠な要素である(提案4)。そこで得られるワークフローの仕組みや研究データそのものが、分野を超えた研究ネットワークの中で発見され、取得され、相互運用性が担保された上で再利用できることが、次世代のオープンサイエンスを切り開く鍵となる(提案5)。その際に、各種データの保護に関する法整備とも連動して、国際的な観点でデータ共有に基づく共同研究を加速させていく必要がある(提案6)。以上のことを、審議依頼3に対する日本学術会議としての提案内容として、次章以降でまとめる。

3 審議依頼への回答

(1) 研究データの共有・公開も含めたオープンサイエンスに対する日本学術会議としての考え方

日本学術会議は、第四の科学⁷と呼ばれるデータ駆動型科学が近年広く進展し多様な成果を生み出す中で、その中核ともいえる研究データの共有・公開を推し進めようとするオープンサイエンスの考えに強く賛同するものである。

第三の科学である計算科学は方程式によって記述可能な世界を対象とし、スパコンを駆使することにより科学に大きな進展を与えたものの、地球、身体、言語のような方程式を明らかにすること自体が困難と考えられる複雑な対象も多々存在する。これらの学理の解明に挑戦すべく、データ駆動型科学が提唱され、今日、多くの分野で推進されるに至っている。「研究 DX」なる用語は、研究にデジタル技術を積極的に活用することにより科学の飛躍的進展を目指そうとする多様な取組の総称といえる。その中でも、研究データを科学者コミュニティで積極的に共有し、最大限利活用することを目指すオープンサイエンスは最も重要な取組といえる。データ駆動型科学は学術全体において大きな潮流となりつつあり、小委員会では、先端的なデータ利用を推進する研究者等から集中的に現況と課題の聴取を進めた。例えば、物質材料学ではロボットを導入し、実験の自動化によりデータの取得を著しく効率化するとともに、得られたデータにAIを適用して新たな仮説を生み、途切れることなく実験を可能とすることで新素材の開発速度を飛躍的に加速するなど、科学実験のスタイルが大きく変容しつつある。この手法はARWと呼ばれている。データは科学的発展だけではなく人類の社会生活にも多大な恩恵を与えている。例えば、ウェブ上の膨大な文書を利活用することにより1億超の対訳テキストが整備され、これを基に高度な言語処理技術を駆使して構築された高品質翻訳サービスが世界の多言語コミュニケーションに革命をもたらしている。他の事例として、古典籍に現れるくずし字を大量に収集し、多くの研究者で共同して現代

⁷ 第一の科学は実験・観測科学、第二の科学は理論科学、第三の科学は計算科学、そして、第四の科学がデータ駆動型科学として、科学のパラダイムシフトが表現されている[24]。

の文字に変換することにより、高精度の AI 認識器を構築する研究成果が報告されている。多様な古典籍の解析効率を抜本的に向上させることができ、また、一般市民も、気軽に古典を読めるようになるなど、人文分野においても大量データの共有が極めて有効であることが明らかになりつつある。すなわち、データは人文・社会科学を更に飛躍的に進化させるものとして、オープンサイエンスにおいても大きな注目を集めている。

このようにデータは多様な分野で大きな変革を創出しつつあるが、研究データの利活用には、データの取得、蓄積、検索、メタデータ付与、それらを統合するデータプラットフォームの構築、人材育成等多様な側面があり、各々に内在する課題を丁寧に明らかにし、解決していくことが肝要である。そのためには、大学等と研究機関に加えて、研究助成団体、学会、出版社等のステークホルダーとの調整も不可欠である。法的取扱いが不十分な無形資産としての研究データの扱いに関しては、産学連携においても、新たなルールが必要である。

以上のような日本学術会議としての問題意識の下、審議依頼の各事項に対して、データ駆動型科学やオープンサイエンスを推進すべく、以下でより具体的に提案することとする。

(2) 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策（管理・活用体制の整備方策、人材確保・育成方策など）

① 【提案 1】研究者が容易に利用可能な研究データプラットフォームの構築

2 で整理したように、我が国は共通基盤を NII RDC として提供し、効果的な基盤開発と運用を実施できている点において、海外に比べて優位性がある。データ活用社会創成プラットフォーム「mdx」[25]や人文学・社会科学データインフラストラクチャー構築推進事業[26]に見られるように、NII RDC との連携や利活用を前提としたプロジェクトも見られるようになってきた。文部科学省による新事業においても、分野プラットフォームとの連携や事例の創出を対象とした取組が実施されているところであるが、これを全国の研究基盤を対象に展開していく必要がある。EOSC ポータルには、データ管理、計算資源、ストレージ、データ分析、データ公開等、2022 年 11 月の時点で欧州の 200 以上の研究基盤が登録されており[27]、サービスの利活用やサービス間の相互運用性を高めるための環境構築が進められている。我が国においては、統合イノベーション戦略推進会議が、これからの研究データの管理・利活用のための中核的なプラットフォーム[28]として位置付けている NII RDC の優位性を生かして、その全国展開を強力に推し進めるべきである。主要な研究プロジェクトのみならず、IT 環境の整備に十分な投資が行われていない中小規模の研究プロジェクトや研究分野にも NII RDC を広範に普及

していくことが、オープンサイエンスやデータ駆動型研究の時代において、我が国の研究力の根底を支えることとなる。

異なる研究分野や異なる IT スキルをもつ研究者に対して広範に基盤の利用を進める際には、図 1 に示した現状の比較的シンプルな構造をもつ NII RDC に更なる対応が求められる。図 2 は、EOSC が示す全体のアーキテクチャである[29]。このアーキテクチャでは、共通基盤としての EOSC Core 及び各分野や機関との連携を実現する EOSC Exchange のレイヤーを定義するとともに、各レイヤーの継ぎ目ない協調を実現するために必要となる支援活動（EOSC Support Activities）と相互運用フレームワーク（EOSC Interoperability Framework）を左右に配置している。このように、EOSC はデータを扱う種々の手続を分割可能な最小限まで落とし込むことで、相互運用性の実現を目指すアーキテクチャとなっている。欧州ではこのアーキテクチャ実現に 2021～2027 年で約 10 億ユーロの予算を組み、サービスの構築と普及を進めている。

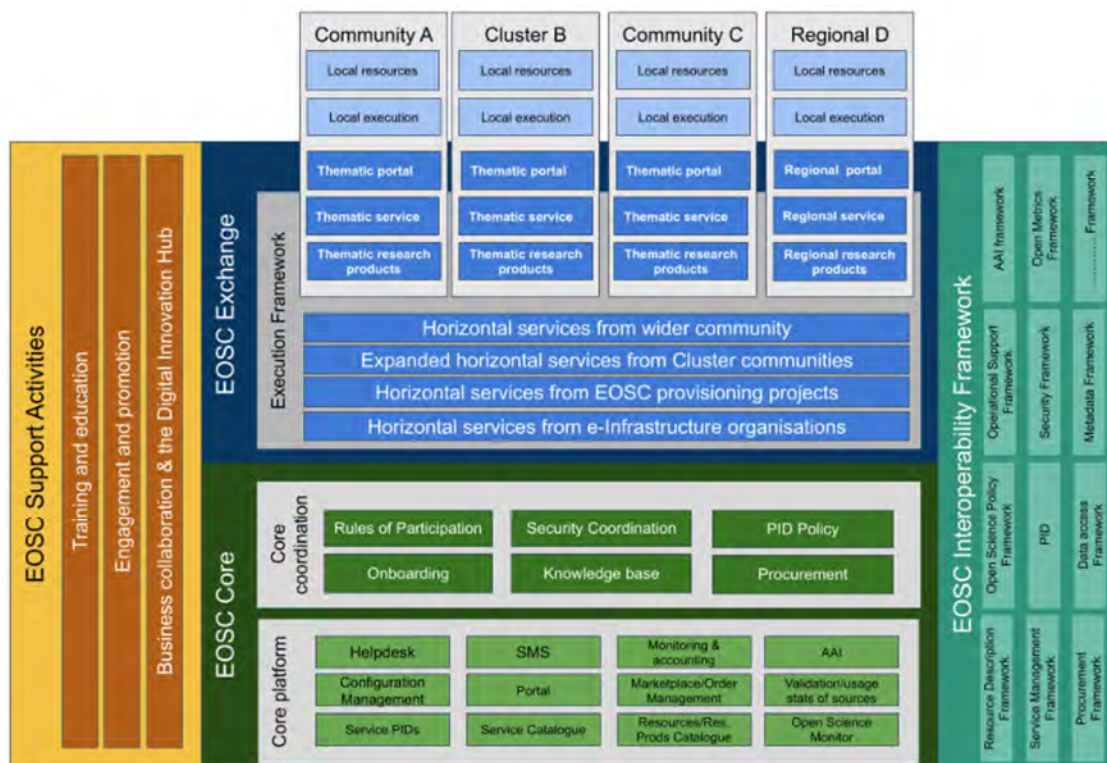


図 2 欧州 EOSC のアーキテクチャ

(出典) Governance, Architectures and Business Models for Data and Cloud Federations: the EOSC and GAIA-X Case Studies (1.0) [29]

EOSC のアーキテクチャは、共通のデータ基盤を提供することに加えて、各分野や研究機関で構築された個別の基盤やサービスを、積極的に連携させることを主眼としているものである。我が国においても、このようなアーキテクチャを参考に、従来までの段階的な取組を国レベル、研究施設や研究分野レベル等の複数の粒度で統合し、普及していく必要がある。そのために多面的な標準やガイドライン等を構築し、NII RDC だけではなく、既存の研究基盤との接続を全国レベルで広げていくことが重要である。国内のあらゆるデータ管理、計算資源、ストレージデータ分析、データ公開のための研究基盤を有機的に連携させていかなければならない。より包括的な視野から国家的な基盤の構築を実施することで、大学や研究機関にとって必要な共通の研究基盤の整備が進み、国際的なデータ駆動型科学やオープンサイエンスの潮流とも歩調を合わせた研究の実践につながっていくことが期待される。

② 【提案2】データプロフェッショナルの育成と多面的な研究評価の実現

データ駆動型科学をあらゆる分野において推進するためには、大量のデータを自在に扱えるスキルや環境、さらには高品質なデータとメタデータの管理が重要となる。データとメタデータを管理・保存・伝送・解析・可視化・共有・公開し、再利用可能にするための一連の処理は、研究者自身が行う場合もあれば、リサーチエンジニア、データサイエンティスト、データキュレータ、データライブラリアンといった専門職（以下、「データプロフェッショナル」という。）によって行われる場合もある。英国や豪州の大学では、機関戦略としてデータプロフェッショナルを全学に配備し、機関の研究力強化につなげている[30][31]。また、データプロフェッショナルが、リサーチソフトウェアエンジニア学会[32]やデータキュレーションネットワークの設立を通じてコミュニティを形成し、情報交換や知見の蓄積を通じて、データ駆動型科学のための技術を機関横断的に進化させている例も見られる（参考調査報告3-2参照）。

データプロフェッショナルは、あらゆる大学や研究機関等に配備できるに越したことはないが、実際には、人材資源や予算等の少ない大学等においては、配備されることは困難である。そのような人材資源や予算の少ない大学等においてもデータ駆動型科学が推進されるよう、データプロフェッショナルがプールされ全国の大学等に派遣される仕組みや、特定の大学のデータプロフェッショナルが他大学等の研究プロジェクトに協力する仕組みなど、データプロフェッショナルが全国の大学等に協力できるような共助の仕組みが構築されることが望ましい。この場合には、異なる機関に所属する研究者やデータプロフェッショナルが連携するチームサイエンスが可能となるように、大学等における研究データに係る方針や制度、責任体制等の整備を通じた、機関の研究データ管理・統治体制（以下「研究データガバナンス」という。）が確立している必要がある。多機関連携においては、特にそれらの

機関が産学官や国内外の大学等、制度枠組みの異なる機関間の連携の場合、機関ごとに研究データに対する権限や責任体制、堅牢な基盤とその運営体制等が明確になっていないと、安心して連携できないからである。このため、海外の大学ではその研究データポリシーにおいて、研究データの取扱いの方針や体制等について明確に定め、機関の研究データガバナンスの基礎を形成している（参考調査報告3-1参照）。現在、我が国の大学において、策定されつつある研究データポリシーは、研究データの取扱いに係る理念的側面を中心に定められているが、より具体的で実効性のあるポリシーの策定が望まれる。

同時に、データプロフェッショナル、あるいは研究者が研究チームにおいてデータの取扱いに特化した役割分担を担うためには、その研究貢献が適正に評価される必要がある。従来からの論文業績を中心とした研究評価体系では、論文業績につながらない活動は敬遠されるからである。また、そうした研究評価体系の下では、研究者間の協力が阻害され、スケールの大きな研究が育たない傾向が生まれる。オープンサイエンスを推進し、複雑な研究課題や社会的課題の解決を図るためには、複数の分野の研究者やデータプロフェッショナル、その他の研究プロジェクト運営に関わる人材等の協力に基づくチームサイエンスの実現が不可欠である。

欧州においては、このような課題意識から2(2)及び参考調査報告1-5に記したような、大掛かりな研究評価改革に乗り出している。欧州委員会や欧州大学協会等が40か国350機関以上の関心表明を得ながら2022年7月に発表した「研究評価の改革に関する合意」の最終版[15]には、表2に示す評価の観点の原則が挙げられ、CoARA[16]が学術システムに係る各種ステークホルダーや各国機関の研究評価改革の取組の情報共有や取りまとめを行いながら、研究評価改革を促進することとなっている。

表2 「研究評価の改革に関する合意」最終版の評価の観点の原則

<p>1) 研究の質とインパクトの評価</p> <p>2) 多様性、包摂性、協働性の評価</p> <ul style="list-style-type: none"> ● 多様な研究貢献 (査読、メンタリング、指導、リーダーシップ、産学・社会連携、アウトリーチ、起業、知識移転等) ● 多様な研究成果 (論文のみならず、データ、ソフトウェア、モデル、メソッド、理論、アルゴリズム、ワークフロー、展示、戦略、政策貢献等) ● オープンサイエンス実践 (情報・データ共有、学術界内のオープンコラボレーション、社会

のアクターとの連携)

※ 個人としてだけでなく、チームや協働性の単位からも評価することに留意

(出典) European Commission et.al.: “Agreement on Reforming Research Assessment”, 2022[15]

日本学術会議科学者委員会研究評価分科会が2021年11月に発出した提言「学術の振興に寄与する研究評価を目指して一望ましい研究評価に向けた課題と展望」[33]では、「(前略) 定性的な研究評価を原則とし、定量的評価指標を補助的に活用することが望まれる。」と述べている。また、文部科学省において2021年に取りまとめられた政策評価から俯瞰するオープンサイエンス時代の研究評価の論点検討会「新しい時代を見据えた研究開発評価の論点—よりよい研究活動の推進のために—」[34]も、表2に示すこれらの研究評価の観点と問題意識において軌を一にしていると考えられる。ただし、日本では研究評価の観点が議論され始めたばかりで、研究評価の見直しの具体化に向けた議論はまだ開始されていない。

なお、データを含めた多様な研究貢献や成果を研究評価に組み込む際には、それらのレベルについても留意する必要がある。例えば、研究を駆動するデータと、論文に対応する知識生産の単位としてのデータとでは意味合いが異なることから、議論が必要である。また、論文の「査読」に相当するデータの妥当性の保証についても検討が必要である。

データ駆動型科学とオープンサイエンスという潮流の中で日本の研究が世界的に評価され、ひいては海外の研究者も日本の研究者とのチームサイエンスを望むような研究環境となるように、日本においても研究評価と学術システムの更なる検討が早急に求められる。

③ 【提案3】 モニタリング機構に基づくデータ駆動型研究の不断の改善

オープンサイエンスは、その実現のための体系的かつ長期的な戦略的投資を必要とする一方、恩恵も大きい。このことを示すユネスコ⁸のオープンサイエンス勧告では、オープンサイエンスを構成する4本柱の一つとして「知識基盤」を挙げている[35]。例えば、NII RDCの検索基盤であるCiNii Researchには、国内の様々な学術成果が集約され、いわゆる知識基盤(Knowledgebase: KB)が形成されつつある。

しかしながら、既に情報流通網が確立された論文や本に関する情報とは異なり、研究データに係る情報は依然として十分に集約されず散在している。

⁸ ユネスコ(国際連合教育科学文化機関、United Nations Educational, Scientific and Cultural Organization U.N.E.S.C.O.)は、諸国民の教育、科学、文化の協力と交流を通じて、国際平和と人類の福祉の促進を目的とした国際連合の専門機関。

効率的な知識発見やデータ駆動型研究のためにはそれぞれの分野や研究機関ごとに管理・公開されている情報を統一的に集約できる仕組みが不可欠である。そのために国内に散在する知識基盤を API⁹やデータの相互運用によって連携させるハブとして機能し、最適化された分散型データフローの中でより迅速に結果を得るためのフレームワークが必要となる。

欧州の EOSC では Interoperability Framework[36]を定義することにより、域内の全ての知識基盤の情報を集約する仕組みを整えつつある。FAIR4Fusion¹⁰等のいくつかのプロジェクトでは、分野ごとの実験施設までも情報流通経路に結合させ、統合化させる試みがなされている[37]。それらを総合的に知識基盤へと昇華させるのが EOSC における検索基盤 OpenAIRE[38]の役割である。OpenAIRE では 600 を超える日本の機関リポジトリ[39]も含めた世界中の学術情報を取り込み、OpenAIRE Research Graph と呼ばれるデータの関係性を軸にした知識基盤を構築している。

欧州では、さらに 2027 年まで実施される Horizon Europe の第 9 期プログラムに基づきオープンサイエンスへの対応[40]や基盤の品質に対しても評価体制を確立させる予定としている。評価軸としては、大きく、科学的な影響、社会的影響、経済的影響の程度を設定し、このうち科学的な影響の 1 項目として、知識の普及とオープン化の促進を掲示している[41]。欧州では、このような評価を実施しつつ、次世代の評価軸を発見してさらに反映させていくライフサイクルを設定しており、組織や研究者の成果に関する統計情報の抽出や比較を円滑に実施できるためのモニタリング機能の開発や運用を担う機構の整備が進められている。モニタリング機構の整備としては、このほかにも英国において Jisc Monitor UK[42]、researchfish[43]、米国において SHARE[44]、CHORUS[45]、南米において LA Referencia[46]等の取組が行われている。

上記のとおり、海外で急速に進む知識基盤の形成及びその高度化を踏まえると、日本国内でも各研究拠点や情報基盤間の情報流通フローを恒常的にモニタリングするために必要な機能の実装と、それを運用する環境を整備し、国内の学術情報・知識の需要と生成の動向を網羅的に観測することで、学術情報流通の需要と供給の結び付きを強化するべきである。これにより学術研究機関の研究者だけでなく、企業や市民も利用することができる知識情報の体系化と利便性が実現できる。このようにして実現される全分野統合型の知識基盤の内部では、格納された学術情報が自動的に統合され、知識や成果情報として体系的に処理される。これにより出版された成果情報の引用関係や、人員が引き継がれた研究計画同士の連鎖情報、公的資金の助成を受けたプロ

⁹ ウェブサービス等のソフトウェアが互いに情報をやり取りするために用意されるインターフェースの仕様。

¹⁰ 2019 年に開始した、欧州の核融合分野における FAIR 原則の適用と複数の研究機器の相互運用を実証するためのプロジェクト。

ジェクトに紐づく成果や参画研究者、研究機関ごとの計画立案の傾向等、体系化されたデータを利用することで、豊かな相互関係をもつ情報が導かれる。

さらに、従来は論文の引用関係の分析に基づいて、研究力の評価やイノベーション創造に向けて取り組むべき課題の把握が行われてきたが、今後は研究データを含むあらゆる学術情報の生成・流通の統計から研究の初期段階の萌芽も見逃さず把握可能な知識基盤環境を整備して、研究動向・研究力の分析を強化すべきである。そのような取組は、学術研究の再現性確保や次世代への継承に加え、分野を超越して隠された関係性を持つデータの発見や学際研究の糸口を掴むことができ、研究評価や傾向分析等の研究管理に対しても有用である。例えば OpenAIRE では OpenAIRE Research Graph において収集した研究データを内部処理し、共同研究の傾向情報や、特許や社会への貢献、企業情報と合わせた評価分析を実施して、OpenAIRE Monitor [38] から機関担当者や研究者に情報を還元している。公開され相互運用可能となった情報が次の研究活動の切欠や社会の利益として機能し、オープンサイエンスによる新しい知の循環を形成することができる基盤を構築すべきであり、研究データの他者との共有がどの程度進められているか、それによって成果が拡大しているか、データに関連するツールはどの程度利用されているか、データはどの程度容易に検索できているかなど、データ駆動型研究環境を可観測化し、不具合を常に監視するモニタリング機能の整備が肝要である。

(3) 各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために考慮すべき事項（研究者間の連携、情報技術や計算資源の活用事例など）、データ共有への具体的取組方策（データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など）

① 【提案 4】 研究自動化 (ARW) に向けた情報技術、計算資源の集約

科学発見のプロセスに関し、実験自動化によって試料作製や実験条件を AI が探して提案することにより、実験時間を短縮しつつ、かつ多くの可能性の中から最適解を探すアプローチが研究成果として報告されている [47]。米国における 4 つの学術機関から成る The National Academies of Sciences, Engineering, and Medicine は、2019 年に科学全域に共通する ARW に関する調査委員会を設置し、情報技術の動向、採用に当たっての勧告、それを実現するために必要な共通性とドメインユニーク性について整理した報告書を 2022 年に出版した [2]。報告書では、科学発見の加速化に加え、人的資源、システムの持続性、研究者へのインセンティブ評価、データに関するプライバシー保護等の網羅的な観点から、ARW を推進する重要性について述べられている。材料科学、素粒子物理学、創薬、生化学、天文学、気候科学、人文情報学、社会科学、行動科学等の様々な分野において ARW の実践は進められており、この新しい研究アプローチにより、科学の在り方が大き

く変わろうとしている。小委員会においても、材料科学における ARW の実践について最新動向の話題提供を受けた（付録 1、4 参照）

図 3 は、ARW の研究プロセスをワークフローとして捉え、その構成要素と関係性を示したものである。研究ワークフローの自動化（ARW）により、実験結果の追跡、データの記録、出所の確立及び信頼性が自然と確保される。研究におけるプロセスやステップを自動化して記録することで、研究に対する相互運用性、透明性、再現性、複製性、及び研究者の責任感を高めることができる。国際的かつ学際的な研究や追試のための共同作業にも大きく貢献する。このような ARW の特質は、オープンサイエンスの目指す科学の在り方と非常に親和性が高い。ARW の原則としても、利用するシステムやツール、プラットフォームはオープン性が担保されていることが重要視されている。こうした ARW の考え方が広がっていくことにより、単一ラボの中で研究の自動化を実現するだけでなく、異なる機関が保有するコアファシリティをつなぐ地球規模での研究基盤の構築へと広がっていく。

国レベルで構築される研究基盤においても、ARW の思想に基づく仕様が実装されていくことが重要である。データ駆動型科学が ARW により画期的な時代を迎えるためには、研究基盤の機能拡張と同期させることが不可欠である。若手研究者も含めたできるだけ多くの研究者が ARW による新しい科学の実践に参画できるように、我が国における戦略的な科学技術政策の一つとして、その具体策の検討を早急に進める必要がある。

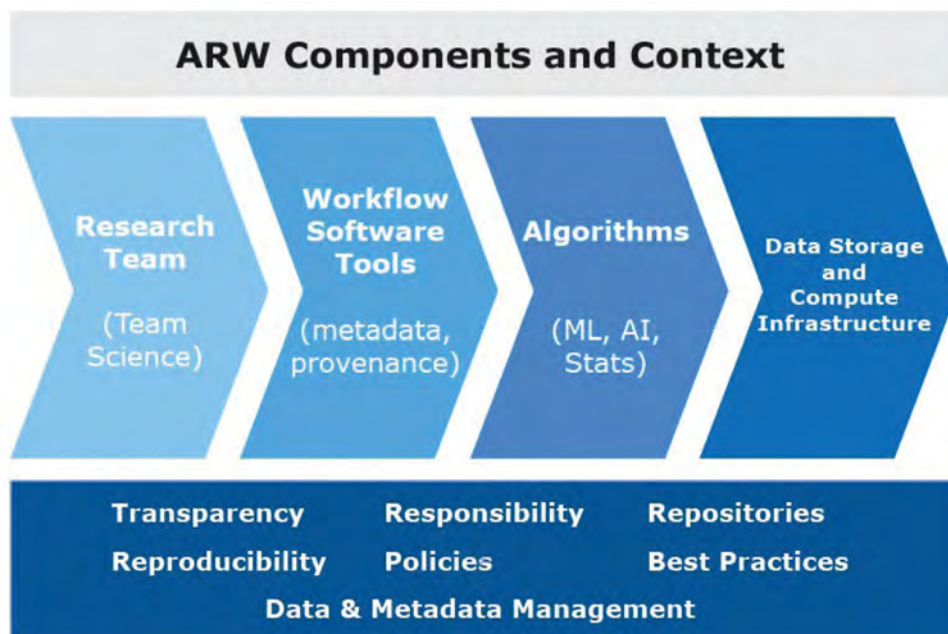


図 3 研究プロセスをワークフローとしてとらえ自動化する ARW の構成要素と関係性

(出典) Automated Research Workflows for Accelerate Discovery: Closing the

② 【提案5】 分野を越えた連携を実現する FAIR 原則の追求

近年の学術研究は必ずしも分野ごとの枠組みで語られることは多くなく、分野横断的（学際的）な取組や連携の下、論文だけでなく研究データも介した幅広い視野で実施されている[48]。学術分野の垣根を越えた連携が、新しい知見の創出や学術研究の高度な発展、ARW の加速に結び付くことは既に自明であり、推進されるべき方向である。一方、各学術分野の内包する情報は多様で、別分野の知識や考え方を結び付けるには相当の努力を必要とする。このような異分野の情報、特にメタデータを共通語彙で表現し研究データに付与するためには、データ利活用のために必要とされる情報を認識し、分野ごとに標準化を行う必要がある。

研究データの適切な共有、公開を行う基準として、データ駆動型科学を推進させる学術コミュニティの慣習である「FAIR 原則」がある[49]。FAIR とは「Findable、Accessible、Interoperable、Reusable」の頭文字を表現したものであり、表3に示すとおり研究データの公開を実現するための理念が、計15項目の原則として定められている[50]。

表3 データ共有の基準としての FAIR 原則

To be Findable: (見つけられるために)	
F1.	(メタ) データが、グローバルに一意で永続的な識別子 (ID) を有すること。
F2.	データがメタデータによって十分に記述されていること。
F3.	(メタ) データが検索可能なリソースとして、登録もしくはインデックス化されていること。
F4.	メタデータが、データの識別子 (ID) を明記していること。
To be Accessible: (アクセスできるために)	
A1.	標準化された通信プロトコルを使って、(メタ) データを識別子 (ID) により入手できること。
A1.1	そのプロトコルは公開されており、無料で、実装に制限が無いこと。
A1.2	そのプロトコルは必要な場合は、認証や権限付与の方法を提供できること。
A2.	データが利用不可能となったとしても、メタデータにはアクセスできること。

<p>To be Interoperable: (相互運用できるために)</p> <p>I1. (メタ) データの知識表現のため、形式が定まっていて、到達可能であり、共有されていて、広く適用可能な記述言語を使うこと。</p> <p>I2. (メタ) データが FAIR 原則に従う語彙を使っていること。</p> <p>I3. (メタ) データは、他の (メタ) データへの特定可能な参照情報を含んでいること。</p> <p>To be Re-usable: (再利用できるために)</p> <p>R1. メタ (データ) が、正確な関連属性を豊富に持つこと。</p> <p>R1.1 (メタ) データが、明確でアクセス可能なデータ利用ライセンスと共に公開されていること。</p> <p>R1.2 (メタ) データが、その来歴と繋がっていること。</p> <p>R1.3 (メタ) データが、分野ごとのコミュニティの標準を満たすこと。</p>
--

(出典) FORCE11: The FAIR Data Principles, 2016. NBDC 研究チーム(訳) [51]

各分野のリポジトリや国際団体においても FAIR 原則は標準として認識されており、生命科学の国際蛋白質データベースとして著名な UniProt では早くから FAIR 原則対応に注力し、その啓蒙も進めていた [52]。世界的なメガファーマが参画する国際非営利組織 Pistoia Alliance においても、製薬に関する情報の FAIR 化のためのプロジェクトが設置されている [53]。

これらの分野における取組を受け、欧州では、第 8 期研究開発ファンディングプログラム (FP8) である Horizon2020 において、FAIR 原則を採用した [54]。並行して、研究データが FAIR に準拠していない場合のコスト分析が行われ、少なくとも年間 102 億ユーロの経済損失が生じるとの試算がなされている [55]。これらを背景に、EOSC のリポジトリを推進し FAIR の認証を提案するプロジェクトである FAIRsFAIR [56] や、分野ごとのケーススタディを収集し網羅的な実施を推進する WorldFAIR プロジェクト [57]、人文社会科学系分野の FAIR 化を進める CO-OPERAS [58] 等が活動しており、FAIR 原則を全体の基準とした分野横断的なデータ共有の実践が進んでいる。また、米国の研究所であり助成機関でもある国立衛生研究所 NIH でも、FAIR 原則がデータサイエンス戦略のために取り入れられている [59]。さらに、分野間の基準の差異による摩擦をさらに減らす取組として、2022 年には FAIR Digital Objects [60] のような考え方も議論されている。

日本国内では官民共に分野間連携の軸となる FAIR 原則等の国際標準を軸とした議論がこれまで十分ではない。国内に散在するデータの共有についても、FAIR 原則に基づく経験の共有を行うための土台となる取組がまだ存在

しない。少なくとも公的助成資金によって創出された研究データについては、FAIR 原則に基づく国際的なフレームワークや一定の標準化された体系の下で管理することが肝要であり、そうしたデータを共通基盤である NII RDC に格納し流通させることも必要である。そのためには、分野ごとに FAIR 原則を追求して分野間連携を実践するだけでなく、日本学術会議や学会等において踏み込んだ議論がなされることが望ましい。

③ 【提案 6】法制度面でのデータガバナンスの構築

データ駆動型科学を進める上では、データの保護と利用に関する法制度面でのガバナンス体制を構築していく必要がある。個人情報については、2021 年の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。）の改正に基づき、従来、個人情報保護法上の各種義務の適用除外とされてきた学術研究機関による学術研究目的での個人情報の取扱いについて、新たに安全管理措置等の規律が設けられることになり、学術研究機関としての自主規範の策定・公表等が求められている。同改正については、2022 年 7 月に NII が「オープンサイエンスのためのデータ管理基盤ハンドブック-学術研究者のための“個人情報”の取扱い方について-」[11]を公表したところであるが、各学術研究機関及び学術界全体として、今後も法令遵守に留まらない、データ活用に関する倫理教育や共通認識の醸成を含む自主的な取組を進めていく必要がある。

同改正における学術研究機関に関する新たな規律は、2019 年に行われた EU による我が国への GDPR¹¹ 十分性認定を、学術研究分野にも適用可能としていくことを視野に入れて設けられたものである。G7 等の枠組みで DFFT (Data Free Flow with Trust) に関する議論が進む中、諸外国とのデータ共有に基づく国際共同研究を円滑に実施していくためにも、我が国の学術界として国内外の関係者から信頼を得られる個人情報保護の施策を進めながら、学術研究分野の国際的なデータ共有の枠組みを構築していく必要がある。

昨今では、データの社会的・経済的重要性の高まりに伴い、個人情報に留まらず、産業データ等の非個人情報についても、経済的価値や安全保障の観点も考慮したより高いレベルの管理体制構築が求められつつある。国際的にデータ主権の議論が拡大する中、各種データの管理強化は不可避であり、知的財産権や契約を含めたデータに関する法的権利の整理や、きめ細かなアクセス管理を可能とするデータ基盤の構築、経済安全保障関連法制への対応を行いつ

¹¹ General Data Protection Regulation（一般データ保護規則）の略であり、2016年に制定され、2018年に施行された EEA（European Economic Area）域内の個人データ保護を規定する EU 法。十分な個人データ保護の水準を有すると認められた域外の国や地域への円滑な越境個人データ移転を可能とする十分性認定制度を有しており、我が国は民間部門について 2019 年に十分性認定を受けたが、学術研究機関等が学術研究目的で個人情報を取り扱う場合については、2021 年改正前の個人情報保護法が一律の適用除外としていたことから、現状において同認定の対象に含まれていない。

つ、人類共通の公益に資するという学術研究の使命に鑑みた、可能な限りのオープンなデータ共有を実現していく必要がある。

EU では、現在、データガバナンス法やデータ法案、欧州ヘルスデータスペース法案¹²等の、従来の個人情報保護法制や知的財産保護法制のアプローチとは異なるデータ活用の促進に焦点を当てた法制度の整備が進められており、その中には、企業や公的機関が保有するデータを学術研究に利用可能とするための新たな枠組みも含まれる。我が国においても、学術研究分野の個人・非個人情報の保護に関する施策を進めながら、パンデミックへの対応を加速させるための EHR (Electronic Health Record) データの活用を含め、学術研究においてより多くの研究データの活用を可能とするための新たな法的アプローチを検討する必要がある。

4 おわりに

今回の回答は、現時点においてとりわけ重要と考えられる観点を取り上げたが、更に検討を要する多くの課題が残っている。例えば、産学連携はデータの共有によって加速化できると期待されるが、大学等の研究機関においては、データのオープン化が原則となりつつあるのに対し、企業においては、保有するデータは極めて貴重な財産であり、原則、非公開とされるのが実情であるため、産学連携研究で生じたデータのオープン化は困難である場合が多い。また、秘密計算等のデータ共有に関する技術の研究開発を進める必要もある。さらには、企業にとって、顧客に提供する機器やサービスを介して得られるデータは顧客に帰属するものとして捉えられることが多く、そのようなデータを用いた産学連携研究を行う上では、広く産業界と学術界の協力に基づくガイドライン等の整備が不可欠である。

また、本回答では、研究データの共有・公開も含めたオープンサイエンスを議論の対象としたが、データを処理するプログラムとの一体的管理は不可欠であり、版管理、バグ管理等の多くの課題が依然として残っている。NII RDC ではコード管理も進めている。東北大学の大林茂教授らが進めている流体科学デジタルツインの研究は、まさに、シミュレーションコードとの融合であり、航空機関連の開発と認証の姿を大きく変容させつつある (付録 2)。データやコードに留まらず、計算資源、ストレージ資源、引いてはエネルギー資源の最適化が研究の加速に大きな役割を果たすと考えられ、そのようなシステム全体についての研究も必須である。

人文科学、社会科学、行動科学において大量のデータに基づく ARW による新たな知の創出への期待が高まっている。人に関わるデータの学術利用に関しては、【提案 6】で述べたように、個人情報保護をしながらより多くのデータ

¹² 欧州における健康医療関連データの安全な利活用を促進するための仕組みである EHDS (European Health Data Space) を構築するために、2022 年に欧州委員会により提案された EU 法案。

の活用を可能とする仕組みづくりが必要である。法制度面でのデータガバナンスの構築は重要だが、それだけではこの問題は解決しない。例えば、人間の活動に関わる過去の歴史データを公開する場合、現行の個人情報保護法制ではカバーしきれない複雑な人権問題が生じる可能性がある。また、安易な研究資料の公開は、研究資料の所有者と研究者の信頼関係を毀損する可能性もある。最先端のデータ秘匿化技術等の利用も視野に入れながら、適切なデータ公開のルールを研究者と社会の間で早急に確立することが望まれる。あわせて、オープンサイエンス時代に求められる研究者の責任も新たな形となる。豪州においては、「責任ある研究実践のための行動規範 2018」において「研究におけるデータと情報の管理」が掲げられたことを受けて、大学における研究データガバナンスの構築が進んでおり、また、科学技術・イノベーション政策においても、研究倫理 (ELSI: Ethical, Legal and Social Issues) と責任ある研究・イノベーション (RRI: Responsible Research Innovation) が重要なテーマとして取り組まれている [61]。オープンサイエンスが研究倫理やデータガバナンスに与える影響等について、それ自体が社会学のテーマとなりながら、実践が繰り返される中で教育にも反映されていくものとなる。

データ駆動型科学の活用が見込まれる領域は著しく広く、今後もそれぞれの学術分野における固有の事情が明らかになると予見されるが、少なくとも学術分野においては、日本学術会議が今後も丁寧に課題を把握しつつ、一步一步研究データの取扱いの方向性を明らかにしていくことが望ましい。加えて、保護と利用を適切にバランスさせた研究データの取扱いは、現時点において全世界に共通した課題であり、機動的かつ柔軟なデータ流通を早急に実現することにより、我が国の研究力を著しく高めることは間違いなく、データ駆動型科学の発展を支えるオープンサイエンスの積極的な推進が望まれる。

＜参考調査報告 1＞欧州の状況

1 汎欧州型のオープンサイエンス基盤の構築（EOSC）

オープンサイエンスを推進するための基盤として、国際的に最も注目されているのは、欧州による EOSC である。EOSC は、欧州諸国の既存の研究基盤を連携させ、サービスへの信頼度の高いアクセスや、分野、社会、地理的境界を越えて研究データの共有や再利用を実現することを目的としている [12]。2015 年に最初の構想が発表された EOSC は、2018 年の正式発足から 2020 年までの初期段階を経て、現在は発展段階（2021～2027 年）に位置付けられている [62]。現在の EOSC は、欧州全体の研究開発支援プログラム「Horizon Europe」の下、欧州委員会や関係国との連携により、200 万人の研究者にサービスを提供することを目指している [63]。欧州委員会は、発展段階における EOSC への開発投資について、欧州委員会と EU 加盟 27 か国から合計 9 億 9000 万ユーロを拠出すると発表している [13]。

こうした汎欧州レベルの広域性、投資規模の大きさとともに、EOSC が提供するサービスの包括性も大きな特徴といえる [36]。単なる研究インフラの基盤を提供するだけでなく、それらを有効活用するためのサービス、相互運用するための標準化や人材育成 [36]、産学連携 [64]、さらには法令遵守に至るまで [65] [66]、研究活動に関わるあらゆる領域を包括する「統合環境」の構築を進めているといえる。

2 研究基盤への戦略的な投資（ESFRI）

2002 年に設立された ESFRI は、各国政府、科学界、欧州委員会が集まり、欧州における研究基盤に対する予算配分の審議機関として、一貫した戦略主導型の支援をしている。最新のロードマップ 2021 では、EBRAINS（欧州脳研究インフラ）、EIRENE RI（欧州における環境暴露評価のための研究インフラ）、ET（アインシュタイン望遠鏡）、EuPRAXIA（アプリケーションに優れた欧州プラズマ研究加速器）、GGP（世代とジェンダープログラム）、GUIDE（デジタル欧州における成長）、MARINERG-i（洋上再生可能エネルギー研究インフラ）、OPERAS（学術コミュニケーションを通じた欧州研究圏におけるオープンアクセス）、RESILIENCE（宗教研究インフラ）、SLICES（コンピューティング/通信実験研究向け科学的大規模インフラ）、SoBigData++ RI（ソーシャル・マイニングとビッグデータ分析のための欧州統合インフラ）という 11 件の研究基盤構築のプロジェクトを発表している [14]。これら新規プロジェクトの投資総額は 41 億 6,400 万ユーロ（平均 3 億 8,000 万ユーロ）に達する。2018 年の ESFRI ロードマップでは 6 億 7,400 万ユーロ（平均 1 億 1,200 万ユーロ）、2016 年のロードマップでは 8 億 5,600 万ユーロ（平均 1 億 4,300 万ユーロ）の計画を発表しており、これまでで最高額の新規研究基盤への投資計画となっている [67]。

ESFRI の成果は、EOSC に有機的に統合され、欧州全体としての研究基盤を構

成することが目標とされている。現状では、ESFRI 起源のサービスが、EOSC の提供するサービスのカタログサイト「EOSC Portal」から検索可能となっている程度であるが、今後は、EOSC が提供する相互連携のための標準仕様に則ったサービスの開発が基本となり、欧州全体としての研究基盤の自律的な連携が加速してくるものと考えられる。

3 FAIR 原則に対する具体的なアクション

「Findable、Accessible、Interoperable、Reusable」の頭文字で表現される「FAIR」は、オープンサイエンスの実践に向けての象徴的な標語として、様々なプロジェクトでその達成基準に関する検討が進められてきた[50]。そこで明らかになってきた分野間でのばらつきや、多様化した解釈の摩擦や複雑さを解決するために、最近では分野別研究データの標準の摺合せや、全体を俯瞰するカタログ化等が進められている[68]。FIPs (FAIR Implementation Profiles) と呼ばれる方法論により、分野別の研究データの FAIR 性を体系化するという取組もある[69]。多様化した FAIR 原則への解釈を収束させるフェーズへと移行しており、EOSC の具体化とともに、幅広い分野の研究者が FAIR の実践に自然と取り組むことができる状況が整いつつある。

4 オープンサイエンスのモニタリング基盤

ユネスコの「オープンサイエンス勧告」は、「加盟国は定量的及び定性的な手法を組み合わせ、オープンサイエンスに関連する政策及びメカニズムをモニタリングする必要がある」と述べている[35]。それと同調する形で、フランス[70]やドイツ[71]等の国々で論文のオープンアクセス、研究データの公開状況等をモニタリングできるサービスの提供が進んでいる。代表的なものとして、OpenAIRE が提供する OpenAIRE Monitor[38]が挙げられる。提供されている機関向けのダッシュボードでは、助成、研究成果、オープンサイエンス、コラボレーション、インパクトという項目に分けて情報が提示される。オープンサイエンスという項目では、論文だけではなくデータセットについても、再利用可能なコンテンツの割合等を確認できる。コラボレーションでは、複数の研究プロジェクトによる研究成果の件数等について経年変化等を見ることができる。機関向けのダッシュボードは提供が始まったばかりであり、その機能は日進月歩で拡張されている。ダッシュボードの背後には OpenAIRE Research Graph[72]という構造化された学術情報があり、これらをオープンデータとして利用に供することで、透明性のある指標の提供を実現している。

5 研究力評価に関する新しい考え方

近年は、欧州を中心として、オープンサイエンスの推進という文脈で研究評価体系の改革が試みられている[73]。オランダでは、キャリアパスの多様化と活性化、オープンサイエンスの促進を目的として、研究者を個人単位のみなら

ずチーム単位で評価することに全国規模で取り組んでいる[74]。欧州委員会におけるオープンサイエンスに関する作業部会では、「オープンサイエンスの実践活動は、全てのレベルの研究者の採用、昇進、研究助成等において、評価されなければならない」という趣旨の提言を公開している[75]。これらの議論を受けて2022年2月には「研究評価改革のためのパリ宣言」[3]が発表され、欧州委員会、助成機関、大学・研究機関、学協会・研究者等が大同団結して研究評価改革についての議論を加速させている。さらに、欧州委員会や欧州大学協会らは同宣言を受けて、「研究評価の改革に関する合意」の最終版を2022年7月に公開した[15]。40か国350機関が、先だって2022年1月に公開された草案に関心表明を示している。この改革を実行するために、2022年12月にはCoARA[16]が正式に設立された。研究の質やインパクトの向上につながる研究評価改革は、「欧州研究圏政策アジェンダ2022-2024」[17]にも位置付けられ、こうした動きの中で、オープンサイエンスがけん引する多様な研究力評価の考え方が、より実践へと向かっている。

＜参考調査報告2＞米国の状況

1 研究データを含む公的資金を得た研究成果の即座オープンアクセス方針

OSTPは2022年8月、公的資金を得て生み出された研究成果について、「即座オープンアクセス」（即座OA）の方針を打ち出した[18][19]。先行して欧州にて推進されている「プランS」（2018.9発表、2021.1発効）[76]に追随した形で、今後、同様の方針が世界に広く波及していくことが想定される[77]。

米国の「即座OA」方針は、論文出版と同時に、論文だけでなく、論文に付随する研究データもOAとすることを求めており、今後、研究データの共有・公開が飛躍的に拡大することが想定される。また、こうした取組は、研究成果の視認性と利活用の拡大を通じて、米国の研究力の更なる向上につながると推測される。

2 大学間で連携したデータキュレーションの実践

米国の一部の大学・研究機関では、機関間でパートナー契約を結び、単一の機関ではカバーしきれない多様な専門分野のデータキュレーションに取り組むデータキュレーションネットワーク（Data Curation Network: DCN。以下「DCN」という。）が構築されている[78]。DCNでは、2016年の立ち上げ以来、実務レベルでの機関横断的な協力体制構築とノウハウの蓄積が進められており、2022年11月現在では主導機関であるミネソタ大学を中心に17機関が協定を結び、取り扱うデータは53分野をカバーするに至っている。2022年4月からは米国国立衛生研究所（以下「NIH」という。）データサイエンス戦略室と合同での教育プログラム提供を開始している[79]。NIHの持つ影響力も相まって今後は米国におけるコミュニティの中心となっていくことが想定される。機関横断的なデータサイエンス支援体制のビジネスモデルを実践、検証して

いる事例として着目に値する。

DCN の取組は、国際的にも波及しつつある。オランダではデルフト工科大学のメンバーが主導する Dutch Data Curation Network が立ち上げられ、米国の事例を参照しながら同国の機関事情に合わせたモデル開発が進められている[80]。さらに、このような各国での実践例を足掛かりに、同ネットワークを国際的な枠組みに拡大するための議論が Research Data Alliance 等で継続的に実施されている[81]。

3 米国のサイバーインフラストラクチャ

米国では、NSF や NIH のプロジェクト資金で構築された多様なサイバーインフラストラクチャ群がデータ駆動型科学を支えている。研究分野ごとや地域ごとのボトムアップの取組が支援される一方、米国全体で研究データ基盤を統合しようとするトップダウンの動きは見られない。

NSF の取組として、国内の高性能計算機と地域分散型ストレージを分野横断的に統合する Open Science Grid[82]や、4つの地域別ハブでビッグデータ研究を支援する National Network of Big Data Regional Innovation Hubs[83]が挙げられる。Center for Open Science は、NSF 他の研究助成金を受けて研究支援サービス Open Science Framework を運営し、2021年に40万人の研究者に利用されている[84]。

NIH の取組として、生物医学研究における巨大なデータの共有を促進する Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability[85]が挙げられる。

この他にも、スタンフォード大学人間中心人工知能研究所が音頭を取って AI 研究における研究資源の共有を推進することを目的とした National Research Cloud[86]等、オープンサイエンス研究基盤を目指すボトムアップな取組は枚挙にいとまがない。

<参考調査報告3> 豪州の状況

1 大学における研究データガバナンスの構築

豪州においては、「責任ある研究実践のための行動規範 2018」において「研究におけるデータと情報の管理」が掲げられたことを受けて、大学における研究データガバナンスの構築が進んでいる[87]。大学の研究データ管理ポリシーには必ずといってよいほど、この行動規範への言及があり、学内の研究データガバナンスの考え方や研究データ管理に係る手続、学内体制が詳細に示されている。

大学における研究データガバナンス体制は一般には研究担当理事によって統括され、研究推進部が学内の研究データ管理に係る体制をコーディネートする。研究データ管理を所掌する研究推進部のチームは「研究インテグリティと研究ガバナンス」と呼ばれることもあり、大学において重要な研究ガバナ

スを担っていることが推測される。情報基盤センターやこれに併設される e リサーチ部門では、研究データの解析ツールを含む研究データ基盤を整備・運用する。大学図書館は、研究データ管理一般及び学内の研究データ管理体制についてウェブサイトにて取りまとめ、定期的に研修プログラムを実施する。

豪州の大学における研究データ管理体制は、深刻な情報流出等の問題に触発されていることもあり、現在、一部の大学では、研究データは機関の提供するストレージにしか保存できないルールとなっている [88]。機関内のデータガバナンスを統括する CDO (chief data officer) が、一般の機関データに加え、研究データの管理についても責任を持つ場合も見られる [89]。各機関における研究データガバナンスへの取組が、一層強化されてきている。

2 データと e リサーチを牽引する豪州研究データコモンズ (ARDC)

豪州研究データコモンズ (Australian Research Data Commons, ARDC。以下「ARDC」という。) は、データに関連する標準やスキル、研究基盤を整備することにより、豪州研究コミュニティと産業界の研究競争力につなげることを目的として、2018 年に設置された [20]。ARDC は、豪州国家データサービス (Australian National Data Service, ANDS。以下「ANDS」という。)、豪州 e リサーチ協働ツールとリソース (National eResearch Collaboration Tools and Resources : Nectar)、研究データサービス (Research Data Services : RDS) を前身とし [90]、豪州研究インフラロードマップ 2016 において研究活動の全般においてデータの取扱いと e リサーチ基盤に依存していることが指摘されたことが設立の発端となっている [91]。

活動の柱として、1) 人材育成とポリシー、2) プラットフォームとソフトウェア、3) データとサービス、4) ストレージと計算機を掲げ、データの有効活用につながる統合環境を構築している [92]。データの相互運用枠組みやデータ保存・伝送方式を例にとっても、数多くの検討グループやプロジェクトが形成されており、それが国内の関係者のネットワーク形成とスキル向上につながっている。ARDC による Institutional Underpinnings プログラムでは、25 大学の協力を得て、大学のポリシーやデータの保持・廃棄期間等の大学における研究データ管理枠組みが 7 項目について作成された [93]。大学という枠を超えたこうした取組が、各大学におけるデータガバナンスの構築の進展につながっている。

ARDC は、NCRIS の一部として位置付けられ、2021 年度の予算は 3910 万豪ドルであった [94]。このうち 6 割強がプロジェクトの実施経費であり、一部は競争的資金として豪州の大学に配分される [94]。これらサービスへの門戸は、産業界にも開かれており、学术界のみならず産学連携の発展にも寄与している [20]。

3 豪州協働型研究基盤戦略の実践（NCRIS）

豪州では、国の研究基盤を、研究競争力だけでなく雇用の創出や健康、食糧の確保にもつながる重要な基盤と位置付けている。豪州研究基盤戦略枠組みに基づき、国の研究基盤を常に、時代の要求にこたえる最新のものとしている[95]。豪州研究基盤戦略枠組みは、5年ごとに策定される豪州研究基盤ロードマップと、これに対応する研究基盤投資計画、戦略を実施につなげる NCRIS プログラムから構成される[96]。NCRIS は 2022 年現在、24 のプロジェクトに助成をしており、2022-2023 年には 2.83 億豪ドル、2018-2029 年に 40 億豪ドルの予算配分を得ることとなっている。NCRIS では研究設備を、「①研究を実施するための施設、設備、リソース、②これを運用するための専門職」と定義しており、研究設備有効活用のための人的支援体制の重要性を前面に出している[21]。

4 大学のコアファシリティ等に対する機関横断的な支援体制

大学のコアファシリティとなる物理的な研究設備の技術職員はデータの取扱いに精通していないことが多いことから、コアファシリティの先進的な活用を進める上では、データの取扱いに係る支援・助言体制を整備することが有効である。豪州では、一部の研究設備について機関横断的にコアファシリティを支援する組織が設けられている。Microscopy Australia では、豪州の大学が保有する顕微鏡類を利用するユーザと研究設備をマッチングするサービスを提供する傍ら、各研究設備のデータの取扱いに係る設備のアップグレードや利用のガイドライン等の整備や助言、各設備への人材派遣も行っている[97]。同様の機関横断的なサービスを提供する研究基盤として、豪州画像化設備 (National Imaging Facility)、豪州核科学技術機関 (Australia's Nuclear Science and Technology Organization: ANSTO)、Astronomy Australia がある[98]。

なお、こうした大学のコアファシリティ等に対する機関横断的な支援体制は、NCRIS に基づき予算措置及び方向付けがされている。

5 大学における研究者の研究データの取扱いに係る支援体制

豪州の一部の大学においては、研究者の研究データの取扱いについて、ハンズオンの支援体制を整備し、研究力強化につなげている。例えば、豪州モナッシュ大学では、40 名ほどのリサーチエンジニアを大学の e リサーチセンターに配備し、2000 以上の学内の研究プロジェクトにおける研究データの取扱いにおいて、問題解決に当たってきた[30]。そのように個別に対応した結果として、データの管理・保存・伝送・解析・可視化・共有・公開等の技術について分野横断的な共通性が見出されたという。豪州内における大学ランキングが、9 位から 2 位へと上昇した原動力となっている[99]。大学における大型計算機センターや共用の研究設備の利用方法について人員を配備し、機関内の研究プ

プロジェクトにおいて助言を提供するとともに、時には人員を一定期間派遣することもある。こうした対応も、NCRIS で示された研究のデジタル対応の重要性の考え方に基づいている。

モナシュ大学は ARDC の前身である ANDS を、オーストラリア国立大学や豪州連邦科学産業研究機構 (Commonwealth Scientific and Industrial Research Organization, CSIRO) とともに 2008 年に生み出した大学であり、早い段階から、データ駆動型科学や研究基盤のデジタル対応の重要性を認識していたと考えられる [100]。

同類の取組は他国でも見られ、例えば英国においては 30 以上の大学がリサーチソフトウェアエンジニアを配備しており、このような取組は欧州を中心に、他国でも広がっている [31][32]。

<参考文献>

- [1] National Research Council: 「Enhancing the Effectiveness of Team Science」, 2015. Available: <https://doi.org/10.17226/19007>
- [2] National Academic of Sciences, Engineering, and Medicine: 「Automated Research Workflows for Accelerate Discovery: Closing the Knowledge Discovery Loop」, The National Academic Press, 2022. Available: <https://doi.org/10.17226/26532>
- [3] Paris Open Science European Conference ウェブサイト: 「Paris Call on Research Assessment」, 2022. Available: <https://osec2022.eu/paris-call/>
- [4] Shimizu, R. et.al.: 「Autonomous materials synthesis by machine learning and robotics」, APL Materials, 2020. Available: <https://doi.org/10.1063/5.0020370>
- [5] Matsuda, S. et.al.: 「Data-driven automated robotic experiments accelerate discovery of multi-component electrolyte for rechargeable Li-O2 batteries」, Cell Rep. Phys. Sci., 2022. Available : <https://doi.org/10.1016/j.xcrp.2022.100832>.
- [6] Kanda, G.N. et.al.: 「Robotic search for optimal cell culture in regenerative medicine」, eLife, 2022. Available: <https://doi.org/10.7554/eLife.77007>
- [7] 日本学術会議オープンサイエンスの取組に関する検討委員会: 提言「オープンイノベーションに資するオープンサイエンスのあり方に関する提言」, 2016. Available: <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-23-t230.pdf>
- [8] 日本学術会議オープンサイエンスの深化と推進に関する検討委員会: 提言「オープンサイエンスの深化と推進に向けて」, 2020. Available: <https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-24-t291-1.pdf>
- [9] 国立情報学研究所ウェブサイト: 「AI等の活用を推進する研究データエコシステム構築事業」, 2022. Available: https://www.nii.ac.jp/creded/nii_ac_jp_creded.html
- [10] 文部科学省ウェブサイト: 「AI等の活用を推進する研究データエコシステム構築事業の公募開始について」, 2022. Available: https://www.mext.go.jp/b_menu/boshu/detail/1421775_00007.htm
- [11] 国立情報学研究所オープンサイエンスのためのデータ管理基盤ハンドブックにかかる検討会: 「オープンサイエンスのためのデータ管理基盤ハンドブック～学術研究者のための“個人情報”の取扱い方について～」, 2022. Available: <https://doi.org/10.20736/0002000318>
- [12] EOSC Portal ウェブサイト: 「EOSC」. Available: <https://eosc->

- portal.eu/about/eosc
- [13] EOSC ウェブサイト: 「Draft of Memorandum of Understanding for the Co-programmed European Partnership on the European Open Science Cloud」, 2021. Available:
https://eosc.eu/sites/default/files/20210215_EOSC_MoU_FinalDraft.pdf
- [14] ESFRI ウェブサイト: 「ROADMAP 2021 Strategy report on Research」, 2011. Available: [Infrastructureshttps://roadmap2021.esfri.eu/](https://roadmap2021.esfri.eu/)
- [15] European Commission et.al.: 「Agreement on Reforming Research Assessment」, 2022. Available:
https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/rra-agreement-2022.pdf
- [16] CoARA ウェブサイト: 「Coalition for Advancing Research Assessment, CoARA」, 2022. Available: <https://coara.eu/>
- [17] European Commission ウェブサイト: 「EUROPEAN RESEARCH AREA POLICY AGENDA: Overview of actions for the period 2022-2024」. Available:
https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_era-policy-agenda-2021.pdf
- [18] White House: 「OSTP Issues Guidance to Make Federally Funded Research Freely Available Without Delay」, 2022. Available:
<https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/>
- [19] Office of Science and Technology Policy: 「Ensuring Free, Immediate, and Equitable Access to Federally Funded Research」, Memorandum for the heads of executive heads and departments, 2022. Available: <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>
- [20] Australian Research Data Commons ウェブサイト: 「About the ARDC」, Available: <https://ardc.edu.au/about-us/who-we-are/>
- [21] Australian Government Department of Education ウェブサイト: 「NCRIS Fact Sheet」. Available: <https://www.education.gov.au/ncris>
- [22] 国立情報学研究所ウェブサイト: 「NII and WACREN Sign a Memorandum of Understanding on R&D of Open Science Infrastructure」, Available: <https://www.nii.ac.jp/en/news/release/2019/0315.html>
- [23] 国立情報学研究所ウェブサイト: 「Myanmar to be first country in

- ASEAN Region to launch free and open national research portal」 , Available: <https://www.nii.ac.jp/en/news/release/2020/0221.html>
- [24] Hey, T. et al.: 「The Fourth Paradigm: Data-Intensive Scientific Discovery」 , Microsoft Research, 2009.
- [25] mdx ウェブサイト : 「mdx とは」 , Available: <https://mdx.jp/about/mdx>
- [26] 日本学術振興会ウェブサイト : 「人文学・社会科学データインフラストラクチャー構築推進事業」 , Available: <https://www.jsps.go.jp/j-di/index.html>
- [27] EOSC Portal ウェブサイト : 「EOSC Portal All Catalogs」 . Available: <https://search.eosc-portal.eu/search/>
- [28] 統合イノベーション戦略推進会議 : 「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. Available: : <https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/sankol.pdf>
- [29] Dietrich, M. et. al.: 「Governance, Architectures and Business Models for Data and Cloud Federations: the EOSC and GAIA-X Case Studies (1.0)」 , Zenodo, 2021. Available: <https://doi.org/10.5281/zenodo.4929021>
- [30] Monash University ウェブサイト ; 「Monash eResearch Centre」 . Available: <https://www.monash.edu/researchinfrastructure/eresearch>
- [31] Woolston, C.: 「Why science needs more research software engineers」 , 2022. Available: <https://www.nature.com/articles/d41586-022-01516-2>
- [32] Society of Research Software Engineering ウェブサイト : 「RSE Groups」 . Available: <https://society-rse.org/community/rse-groups/>
- [33] 日本学術会議科学者委員会研究評価分科会 : 「学術の振興に寄与する研究評価を目指してー望ましい研究評価に向けた課題と展望ー」 , 2021. Available: <https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-25-t312-1.pdf>
- [34] 文部科学省政策評価から俯瞰するオープンサイエンス時代の研究評価の論点検討会 : 「新しい時代を見据えた研究開発評価の論点ーよりよい研究活動の推進のためにー」 . Available: https://www.mext.go.jp/content/20210730-mxt_kanseisk02-000017162_s6.pdf
- [35] UNESCO ウェブサイト : 「UNESCO Recommendation on Open Science」 , 2021. Available: <https://en.unesco.org/science-sustainable-future/open-science/recommendation>
- [36] EOSC Portal ウェブサイト : 「EOSC Architecture and

- Interoperability Framework」, 2021. Available: <https://eosc-portal.eu/sites/default/files/EOSC%20Future-WP3-EOSC%20Architecture%20and%20Interoperability%20Framework-2021-12-22%5B17%5D%5B6%5D-2.pdf>
- [37] FAIR4FUSION ウェブサイト: 「Fair for Fusion - open access for fusion data in Europe」. Available: <https://www.fair4fusion.eu/>
- [38] OpenAIRE ウェブサイト: 「OpenAIRE Monitor」. Available: <https://monitor.openaire.eu/>
- [39] OpenAIRE EXPLORE ウェブサイト: 「country="JP"検索結果」. Available: <https://explore.openaire.eu/search/content-providers?country=%22JP%22>
- [40] European Commission ウェブサイト: 「Horizon Europe, open science : early knowledge and data sharing, and open collaboration」, 2021. Available: <https://data.europa.eu/doi/10.2777/18252>
- [41] European Commission ウェブサイト: 「PAVING THE PATHWAYS TO IMPACT IN HORIZON EUROPE」. Available: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/other/events/20210421/impact-following-kip_en.pptx
- [42] Jisc ウェブサイト: 「Monitor UK」. Available: <https://www.jisc.ac.uk/monitor-uk>
- [43] researchfish ウェブサイト: 「Track research and evidence impact with Researchfish by Interfolio」. Available: <https://researchfish.com/>
- [44] Hudson-Vitale, C. et. al.: 「SHARE: Community-focused Infrastructure and a Public Goods, Scholarly Database to Advance Access to Research」, D-Lib Magazine, Volume 23, Number 5/6, 2017. Available: <https://doi.org/10.1045/may2017-vitale>
- [45] CHORUS ウェブサイト: 「CHORUS Institution Dashboard Service」. Available: <https://www.chorusaccess.org/resources/chorus-for-institutions/chorus-institution-dashboard-service/>
- [46] LA Referencia ウェブサイト 「HOME」. Available: <https://www.lareferencia.info/en/>
- [47] Service, R. F.: 「AIs Direct Search for Materials Breakthroughs」, Science, 366, 2019. Available: <https://doi.org/10.1126/science.366.6471.1295>
- [48] 片岡龍峰: 「オーロラと人間社会の過去・現在・未来」, 2016. Available: <http://id.nii.ac.jp/1291/00011002/>
- [49] Wilkinson, M.D. et. al.: 「The FAIR Guiding Principles for

- scientific data management and stewardship」. Sci Data., 2016.
Available: <https://doi.org/10.1038/sdata.2016.18>
- [50] Jacobsen, A. et. al.: 「FAIR Principles: Interpretations and Implementation Considerations」 Data Intelligence, 2 (1-2), pp.10-29, 2020. Available: https://doi.org/10.1162/dint_r_00024
- [51] FORCE11 のウェブサイト: 「THE FAIR DATA PRINCIPLES」, 2016.
Available: <https://www.forcell.org/group/fairgroup/fairprinciples>,
NBDC 研究チーム(訳): 「FAIR 原則 (「THE FAIR DATA PRINCIPLES」 和訳)」, 2019. Available: <https://doi.org/10.18908/a.2019112601>
- [52] Uniprot ウェブサイト: 「Being FAIR at UniProt」, 2016. Available: <http://insideuniprot.blogspot.com/2016/11/being-fair-at-uniprot.html>
- [53] Pistoia Alliance ウェブサイト: 「FAIR Implementation Supported by a FAIR Toolkit for the Life Science Industry」. Available: <https://www.pistoiaalliance.org/projects/current-projects/fair-implementation/>
- [54] European Commission ウェブサイト: 「H2020 programme guidelines on FAIR data management in horizon 2020」, 2016. Available: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- [55] European Commission ウェブサイト: 「Cost-benefit analysis for FAIR research data : cost of not having FAIR research data」, 2019. Available: <https://data.europa.eu/doi/10.2777/02999>
- [56] FAIRsFAIR ウェブサイト: 「The Project」. Available: <https://fairsfair.eu/the-project>
- [57] WorldFAIR ウェブサイト: 「Global cooperation on FAIR data policy and practice」. Available: <https://worldfair-project.eu/>
- [58] OPERAS ウェブサイト: 「CO-OPERAS」. Available: <https://operas-eu.org/international-partnerships/co-operas/>
- [59] NIH ウェブサイト: 「NIH STRATEGIC PLAN FOR DATA SCIENCE」, 2018. Available: https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf
- [60] 1st International Conference on FAIR Digital Objects ウェブサイト: 「Programme themes」, 2022. Available: <https://www.fdo2022.org/programme/programme-themes>
- [61] 科学技術振興機構: 「ELSI から RRI への展開から考える科学技術・イノベーションの変革 —政策・ファンディング・研究開発の横断的取り組み

- みの強化に向けてー」, 2022. Available:
<https://www.jst.go.jp/crds/pdf/2021/RR/CRDS-FY2021-RR-07.pdf>
- [62] European Commission ウェブサイト: 「Directorate-General for Research and Innovation, Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC)」, 2022. Available:
<https://data.europa.eu/doi/10.2777/935288>
- [63] European Commission ウェブサイト: 「Draft proposal for a European Partnership under Horizon Europe European Open Science Cloud (EOSC) Partnership」, Ver.28, May 2020. Available:
https://ec.europa.eu/info/sites/default/files/research_and_innovation/funding/documents/ec_rtd_he-partnership-open-science-cloud-eosc.pdf
- [64] Dietrich, M. et. al.: 「Governance, Architectures and Business Models for Data and Cloud Federations: the EOSC and GAIA-X Case Studies (1.0)」, Zenodo, 2021. Available:
<https://doi.org/10.5281/zenodo.5081865>
- [65] Foggetti, N. et. al.: 「Legal and Policy Framework and Federation Blueprint」, Zenodo, 2021. Available:
<https://doi.org/10.5281/zenodo.5647948>
- [66] EOSC-Pillar ウェブサイト: 「EOSC-Pillar Legal Compliance Guidelines for Researchers: a Checklist」, 2022. Available:
<https://www.eosc-pillar.eu/news/eosc-pillar-legal-compliance-guidelines-researchers-checklist>
- [67] ESFRI ウェブサイト: 「ESFRI announces new RIs for Roadmap 2021」, 2021. Available: <https://www.esfri.eu/latest-esfri-news/new-ris-roadmap-2021>
- [68] EUDAT ウェブサイト: 「Research Data Management Challenges and Available Solutions」. Available: <https://www.eudat.eu/eudat-conference-2022/programme/research-data-management-challenges-and-solutions>
- [69] Schultes, E. et. al.: 「Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence」, 2020. Available:
https://doi.org/10.1007/978-3-030-65847-2_13
- [70] French Ministry of Higher Education and Research ウェブサイト: 「French Open Science Monitor」. Available:
<https://frenchopensciencemonitor.esr.gouv.fr/>
- [71] Forschungszentrum Jülich ウェブサイト: 「Open Access Monitor」. Available: <https://open-access-monitor.de/>

- [72] OpenAIRE ウェブサイト: 「OpenAIRE Research Graph」 . Available: <https://graph.openaire.eu/>
- [73] 船守美穂: 「オープンサイエンスと 21 世紀に求められる研究評価改革 —日本の研究評価改革の議論に欠けているのは何か?」, 日本高等教育学会 第 25 回大会, 2022. Available: <https://researchmap.jp/funamori/presentations/37053116>
- [74] Recognition & Rewards ウェブサイト: Available: <https://recognitionrewards.nl/>
- [75] Working Group on Rewards under Open Science: 「Evaluation of research careers fully acknowledging Open Science practices: rewards, incentives and/or recognition for researchers practicing Open Science」, 2017. Available: <https://data.europa.eu/doi/10.2777/75255>
- [76] Plan S ウェブサイト: 「Why Plan S」 . Available: <https://www.coalition-s.org/why-plan-s/>
- [77] Lem P: 「Asia tipped to follow US lead with open access mandates」, Times Higher Education, 2022. Available: <https://www.timeshighereducation.com/news/asia-tipped-follow-us-lead-open-access-mandates>
- [78] Data Curation Network ウェブサイト: 「About the DCN」 . Available: <https://datacurationnetwork.org/about-the-dcn/>
- [79] National Institutes of Health ウェブサイト: 「Data Curation Network - Event Series (ODSS, NLM)」 . Available: <https://datascience.nih.gov/data-curation-network-event-series>
- [80] The National Coordination Point Research Data Management ウェブサイト: 「A Dutch Data Curation Network」 . Available: <https://www.lcrdm.nl/en/a-dutch-data-curation-network>
- [81] Research Data Alliance ウェブサイト: 「Data Networks BoF」 . Available: <https://www.rd-alliance.org/data-networks-bof>
- [82] OSG ウェブサイト: 「The OSG Consortium」, Available: <https://osg-htc.org/about/introduction/>
- [83] The Big Data Innovation Hubs ウェブサイト: 「About the Hubs」, Available: <https://bigdatahubs.org/about-the-big-data-innovation-hubs/>
- [84] Center for Open Science ウェブサイト: 「Center for Open Science Impact Report 2021」, 2021. Available: <https://www.cos.io/impact>
- [85] National Institutes of Health ウェブサイト: 「About the STRIDES Initiative」 . Available: <https://datascience.nih.gov/strides>

- [86] Ho, D.E. et. al.: 「Building a National AI Research Resource: A Blueprint for the National Research Cloud」, 2021. Available: https://hai.stanford.edu/sites/default/files/2022-01/HAI_NRCR_v17.pdf
- [87] Australian Government National Health and Medical Research Council ウェブサイト: 「Australian Code for the Responsible Conduct of Research, 2018」, 2018. Available: <https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018>
- [88] 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による、豪州クイーンズランド大学、シドニー大学、シドニー工科大学の2022年10月ヒアリング調査による。
- [89] 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による、豪州シドニー工科大学の2022年10月ヒアリング調査による。
- [90] Barker M, Wilkinson R, Treloar A.: 「The Australian Research Data Commons」, Data Science Journal, 18(1):44, 2019. Available: <http://doi.org/10.5334/dsj-2019-044>
- [91] Australian Government Department of Education, Skills and Employment: 「2016 National Research Infrastructure Roadmap - information on development」, 2016. Available: <https://www.education.gov.au/national-research-infrastructure/resources/2016-national-research-infrastructure-roadmap-information-development>
- [92] ARDC ウェブサイト: 「Our Impact」. Available: <https://ardc.edu.au/about-us/our-impact/>
- [93] ARDC: 「ARDC Institutional Underpinnings Framework (Version 1)」, Zenodo, 2022. Available: <https://doi.org/10.5281/zenodo.6410527>
- [94] The Australian Charities and Not-for-profits Commission ウェブサイト: 「AUSTRALIAN RESEARCH DATA COMMONS LIMITED」. Available: <https://www.acnc.gov.au/charity/charities/eca273f3-f5be-e911-a98a-000d3ad02a61/profile>
- [95] Australian Government Department of Education ウェブサイト: 「National Research Infrastructure」. Available: <https://www.education.gov.au/national-research-infrastructure>
- [96] Australian Government Department of Education ウェブサイト: 「National Research Infrastructure Strategic Framework」. Available: <https://www.education.gov.au/strategic-framework>
- [97] Microscopy Australia ウェブサイト: 「About Microscopy Australia」

- . Available: <https://micro.org.au/>
- [98] ARDC ウェブサイト: 「Understanding the building blocks of reality」, 2022. Available: <https://ardc.edu.au/article/ncris-the-power-behind-australias-science/>
- [99] 元・豪州モナシュ大学 e リサーチセンター長、現・クイーンズランド大学副学長（研究とインフラ担当）Paul Bonnington 氏の、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による 2022 年 10 月ヒアリング調査による。
- [100] University of South Australia ウェブサイト: 「Australian National Data Service」. Available: <https://data.unisa.edu.au/dap/organisation.aspx?organisationid=95939>

<参考資料 1> 審議経過

令和 4 年

- 3 月 23 日 内閣府大臣官房総合政策推進室長及び内閣府科学技術・イノベーション推進事務局統括官から日本学術会議会長宛てに審議依頼
- 4 月 15 日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会（第 25 期・第 7 回）メール審議
・分科会及び小委員会設置について
- 5 月 11 日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会（第 25 期・第 1 回）及び同分科会オープンサイエンス・データ利活用推進小委員会（第 25 期・第 1 回）合同会議
・分科会及び小委員会設置の説明
・分科会及び小委員会の役員の選出と承認
・審議依頼について
・今後の進め方について
- 5 月 31 日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会（第 25 期・第 2 回）及び同分科会オープンサイエンス・データ利活用推進小委員会（第 25 期・第 2 回）合同会議
・分科会幹事の選出と承認
・話題提供：「マテリアルインフォマティクス研究に関して」一杉太郎氏（東京大学）
- 6 月 15 日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第 25 期・第 3 回）
・話題提供：「航空におけるデジタルとグリーン」大林茂氏（東北大学）
- 6 月 28 日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第 25 期・第 4 回）
・話題提供：「オープンサイエンス推進のための推論の標準化と科学のパラダイムシフト」桜田一洋氏（慶應義塾大学）
- 7 月 6 日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第 25 期・第 5 回）
・話題提供：「AI 駆動型科学とその学術・社会への影響」高橋恒一氏（理化学研究所生命機能科学研究センター）

- 7月20日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第25期・第6回）
・話題提供：「農業分野で加速するデータ活用～研究開発や栽培での活用例と課題～」三輪泰史氏（日本総合研究所）
- 9月5日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第25期・第7回）
・話題提供：「DMGMORIにおけるDX化の取り組み」藤嶋誠氏（DMG森精機株式会社）
- 9月9日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第25期・第8回）
・話題提供：「建築分野におけるオープンデータ利用」田辺新一氏（早稲田大学）
- 9月29日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会（第25期・第3回）及び同分科会オープンサイエンス・データ利活用推進小委員会（第25期・第9回）合同会議
・小委員会幹事の選出
・話題提供：上東貴志氏（神戸大学）
・意思の表出の申出書について
- 10月18日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第25期・第10回）
・話題提供：「社会科学分野におけるオープンデータの活用：国際関係・人間の安全保障への適用事例を中心に」阪本拓人氏（東京大学）
- 11月1日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会打合せ
・話題提供：「European Open Science Cloud (EOSC) の動向」小野智弘氏（KDDI 総合研究所）
- 11月16日 オープンサイエンスを推進するデータ基盤とその利活用に関する検討委員会 オープンサイエンス企画分科会オープンサイエンス・データ利活用推進小委員会（第25期・第11回）
・審議依頼の回答案についての審議
- 11月24日 オープンサイエンスを推進するデータ基盤とその利活用に関する

る検討委員会（第9回）

- ・ 役員を選出と承認
- ・ 審議依頼の回答案についての審議

＜参考資料２＞審議依頼公文

府総第104号－2
府科事第344号－2
令和4年3月23日

日本学術会議会長
梶田 隆章 殿

内閣府大臣官房総合政策推進室長
笹 川 武

内閣府科学技術・イノベーション推進事務局統括官
米 田 健 三

研究DXの推進－特にオープンサイエンス、データ利活用推進
の視点から－に関する審議について（依頼）

政府と、科学者コミュニティを代表し、学術に関する各分野の有識者で構成されている貴会議とが、歩調を合わせて社会の大きな問題に取り組んでいくことは、国民の皆様のためになり、また、国際社会における日本のプレゼンスを高めるためにも重要です。

貴会議との未来志向の対話の一環として、今般、政府の講ずる施策について、科学者コミュニティ自身の取組として、現場でどう対応しているのか、また、どう対応していくのかを含め、貴会議から施策推進のための具体的な御意見や御提案をいただきたいと考えています。

政府と科学者コミュニティが歩調を合わせて取り組む必要がある分野横断的なテーマとしては、特にオープンサイエンス、データ利活用推進の視点から、研究DXの推進が挙げられます。

政府としては、研究DXにより、実験・理論・シミュレーションに次ぎ第4の科学ともいべきデータ駆動型科学を振興し、我が国の研究力を強化するため、第6期科学技術・イノベーション基本計画及び「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」(令和3年4月27日統合イノベーション戦略推進会議)に基づき、次のとおり取組の方向性を示してきています。

- ・ 中核的なプラットフォームである研究データ基盤システムの整備
- ・ 大学・国立研究開発法人等におけるデータポリシーの策定
- ・ 公募型の研究資金における研究データ管理・利活用の仕組みの導入 等

また、国際的にも、G7科学技術大臣会合において、オープンサイエンスの推進は継続的に議論されてきており、次に掲げるものが主な課題となっています。

- ・ 国際的な研究データ基盤の相互運用性と持続性
- ・ 研究データ共有のインセンティブ(評価の在り方等)

つきましては、来年日本での開催が見込まれるG7、G7科学技術大臣会合及びGサイエンス会合を見据え、貴会議において、オープンサイエンス分野で日本がリーダーシップを発揮するための方向性やそのためにアカデミアとして取り組むべき事項について、また、政府が示している取組の方向性に対する研究現場の対応状況や新たな課題の有無、更なる改善策等について御提案をいただきたく、下記事項について御検討いただきますようお願いいたします。

記

- 1 これまでの日本学術会議における検討を踏まえ、研究データの共有・公開も含めたオープンサイエンスに対する日本学術会議としての考え方の取りまとめ
- 2 大学・国立研究開発法人等において必要となる研究データ管理・利活用のための課題の整理と具体的方策(管理・活用体制の整備方策、人材確保・育成方策など)
- 3 各分野の多様性を踏まえ、今後のデータ駆動型科学の振興のために考慮すべき事項(研究者間の連携、情報技術や計算資源の活用事例など)、データ共有への具体的取組方策(データ共有へのインセンティブ付与のための方策、分野間連携のためのコミュニケーションの在り方など)

<付録 1 >

令和 4 年 5 月 31 日 第 2 回小委員会話題提供

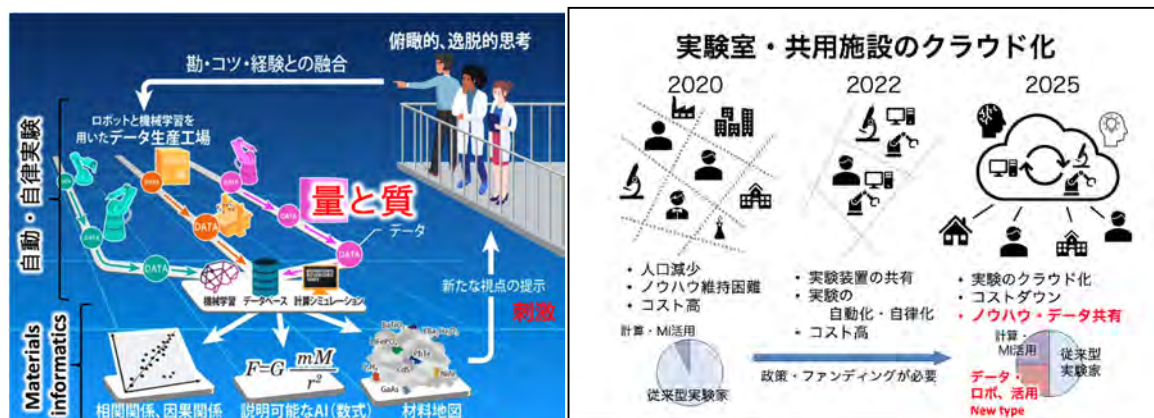
「マテリアルインフォマティクス研究に関して」一杉太郎（東京大学）

発表及び議論の概要

材料科学分野では、マテリアルインフォマティクスと自動・自律実験を組み合わせ、研究者が俯瞰的、逸脱的思考を見出すための仕組み作りが活発化している。実験スループットの向上により探索空間を広げ、セレンデピティにつながることを期待するものである。一杉先生の研究室においては、ロボットを用いた薄膜生成と、過去の実験条件と計測結果から、ベイズ推定により、より良い特性を導く実験条件を予測し次の実験を行うサイクルを自動化した系を既に開発している。実験はリアル空間で行われるためスピードアップには限界があり、今後は研究活動全体の効率化が求められる。そのため、いかに実験室のDX、すなわち実験室のクラウド化を進めるかが一つの鍵となる。クラウド化された実験室では、装置は相互接続され、実験データは自動でクラウドに蓄積される。実験は必ずしも機器のある場所で行う必要はなく、理論研究者がアイデアを直ちに実験化することが可能となり、装置だけでなく、データ、ノウハウ、アイデアが共有される世界となる。

実験室のクラウド化を進めるにはいまだ難しい局面がある。企業は「共有」に対して消極的で、研究者は「勘・コツ・経験を涵養すべし」という意識が強い。実際には勘・コツ・経験も重要であり、全てを共有化することも現実的でない。協調領域と競争領域の区別、適切なデータ管理と具体的なインセンティブを示すことが肝要である。業界では幾つかの動きがあり、例えば JAIMA（日本分析機器工業会）では改竄防止機能を備えた装置データのフォーマット共通化を進めている。

ここで示した構想は、科学一般に適応可能なものであると想定しているが、より広範な材料科学分野を含め全ての分野に適するかは慎重に考慮すべきである。



<付録 2 >

令和 4 年 6 月 15 日 第 3 回小委員会話題提供
 「航空におけるデジタルとグリーン」大林茂（東北大学）
 発表及び議論の概要

航空機設計や運航に関する技術開発において、現在は、流体科学デジタルツインと呼ばれる手法を取り入れる動きが活発化している。大林先生の研究室においても、後方乱気流と空港の離発着に関する研究や新しい材料を使った機体部品開発に関する研究、さらには風洞実験に関する研究において、デジタルツインによるシミュレーションを主体とした研究開発に取り組んでいる。業界では、Certification by Analysis (CbA) と呼ばれる、シミュレーションによる航空機試験の認証を実現するための国際ガイドラインの策定も進められており、航空業界における研究や開発は、データやシミュレーションへの依存度を高めている。しかしながら、産業そのものと密接に関係しているため、その成果を共有する動きはこれからとってよい状況にある。

データの公開や共有が比較的遅れている分野に対し、単純に公開の原則を伝えるだけでは十分ではない。義務のために再利用性の低いデータが公開される事例が山積すると、オープンサイエンスの活動自体に負の側面を与えることになる。東北大学流体科学研究所は、世界最先端の実験設備を有している。この設備で得られる基礎的なデータに関しては、積極的に公開されている。再利用性の高い価値のあるデータが共有されることにより、施設そのものの国際的な評価が高まるという意義が認められているからである。シミュレーションを主体とした他のデータにおいても、その共有や公開への意義を研究者や企業関係者が享受できる状況が、今後構築されていくことが望まれる。



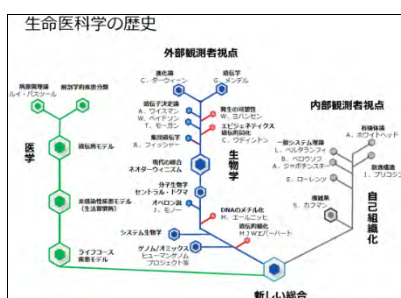
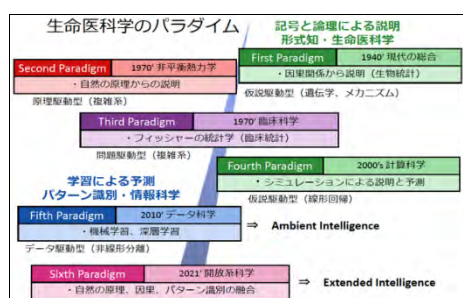
<付録3>

令和4年6月28日 第4回小委員会話題提供

「オープンサイエンス推進のための推論の標準化と
科学のパラダイムシフト」 桜田一洋（慶應義塾大学）

発表及び議論の概要

「生命医科学」に対する形式知、すなわち生命の普遍性に関する知識体系がある一方で、「医療」は病人ごとに個別性が問題となる不可逆なマクロのパターン認識という特徴あり、両者の間には大きな溝があることが知られていた。これをつなげるのが臨床科学であり、オープンサイエンスである。形式知とは記号と論理によって導かれるものとして、知識が蓄積されてきた。一方でAIによる、形式知を使わない方法によって、パターンを識別・学習し予測する情報学が進化し、高い精度の予測等が行えるようになってきた。形式知は、非線形な事象は因果関係からは予測できない点と、自然言語で記述されることの問題があり、データ駆動型科学には、観測境界設定の問題、特徴選択・特徴量の重み付の不完全性の課題がある。こういった異なるパラダイムを統合することが重要になってくる。ではどうやって行うか。近代合理主義に見られる、内部観測者視点と外部観測者視点を結びつけることで、生命科学と情報科学の融合を行うことにつながる。その際のモデルとしては、閉鎖系モデル、平衡解放系モデル、非平衡解放系モデル等のモデルがある。どれを選べばいいとかではなく、どう融合するか、それによって生物の状態をきちんと識別できることが重要で、そのためには、データの融合、状態識別の標準化、状態遷移の標準化、予測信頼性の標準化をやっていく必要がある。信頼性を保証し担保することは簡単ではないが、使いながら信頼性を担保していくことができるようになる。新しい医療の世界として、サイバー空間からフィジカル空間の誘導がなされることになる。ただし、それが行き過ぎた場合、自動的な行動変容を促され、自由な意思決定が奪われる、デジタル専制主義への懸念がある。それを克服するためには、人間中心のデザインが必要である。今後、医療分野のオープンサイエンスを推進する仕組みとして、健康・医療の情報化(DX化)により、データ統合・共有を前提とした、推論の標準化が中核となることだが、非常に難しい問題。ヨーロッパは形式知を重要視していて抜け出せない。そこに日本が新しい知の体系のリーダーシップを取れるチャンスがある。



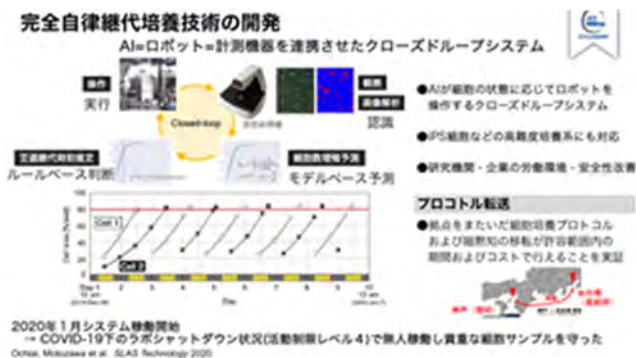
<付録 4 >

令和 4 年 7 月 6 日 第 5 回小委員会話題提供
「AI 駆動型科学とその学術・社会への影響」
高橋恒一（理化学研究所生命機能科学研究センター）
発表及び議論の概要

生命科学分野の研究開発において、AI を科学の道具として研究活動にどう活かせるか、最終的には AI 自体が科学を推進していくことを目指す「AI 駆動型科学」が注目されている。理化学研究所の生命機能科学研究センターでは、AI が顕微鏡画像から細胞の状態を把握して自律的にロボットを制御することで、完全自動の細胞培養技術の開発に取り組んでいる。この技術が実用化されれば、細胞培養の専門家だけが持っている匠の技（暗黙知）を形式知に変換できるだけでなく、実験プロトコルをクラウドアーカイブで公開することで、世界中で即座に（これまで再現が難しかった）生物系実験を再現することができるようになる。

将来的には、大量データの蓄積と AI 技術の高度化によって、仮説→予測→検証という仮説演繹のプロセスと、データ→法則→仮説という帰納のプロセスとを、①ロボットによる実験プロセス自動化プロセス、②AI による仮説空間の高速探索の 2 つの自動化技術を基盤として一つの連続したサイクルが形成され、研究プロセスの統合と自動化が進むことが期待される。

AI 駆動型科学の普及は世の中の研究開発の加速に資するものであることは間違いないが、一方で、日本においては情報漏洩の観点からデジタル化を慎重に考える傾向があるなど、その普及には時間がかかるだろうという指摘もある。形式化された暗黙知はそれ自体が高い価値を持つが、企業が知見をオープンにしてもそれ以上のメリットを見いだせるような仕組みの構築が望まれる。特に、機械産業等アナログ技術を強みとする産業において DX 化を積極的に推し進めるための素地づくり（経営者の意識改革等）が課題である。



第五の科学： AI・ロボットによる研究プロセスの統合と自動化

モデル駆動： 仮説 → 予測 → 検証という仮説演繹のプロセスと
データ駆動： データ → 法則 → 仮説という帰納のプロセスとを、

- 1 ロボットによる実験プロセス自動化
- 2 AIによる仮説空間の高速探索

の2つの自動化技術を基盤として一つの連続したサイクルとして統合してゆく。

目的： 複雑な対象の挙動を説明するモデルの精度を高め、予測と制御の手段を得る。

<付録 5 >

令和 4 年 7 月 20 日 第 6 回小委員会話題提供

「農業分野で加速するデータ活用」三輪泰史（株式会社日本総合研究所） 発表及び議論の概要

三輪先生が担当されてきた SIP 第 1 期、NEDO、NARO 等の農業分野でも DX が推進されているが、実際に農家と連携したスマート農業の取組にはデータの共有・公開とインセンティブ等オープンサイエンスに向けた課題も発生している。農林水産省の農業 DX 構想では、農業従事者は急速に減少しているが、一方で現代では IoT やロボティクスの農地での活用が現実的になってきており、現場では好機会として前向きに捉えている。農林水産省のリードも有りオープンサイエンスは農業 DX の中心になりつつある。農林水産省は農業生産の現場、行政実務そしてデータ基盤の 3 点を中心に農業 DX を行っている。内閣府 SIP 事業では農業版デジタルツインを目指した農業データ連携基盤 WAGRI がある。民間企業、NARO、官公庁から様々な農業データを提供し、農業者や研究者は API から農機メーカー、ICT ベンダーを通じてデータを取得することができる。他事例として、ゲノム選抜 AI や人工気象室を活用した育種法で苗の段階で評価選抜が可能になっており、通常に比べて 2～3 年早く研究結果を得られる。マルチオミクスの相関ネットワーク解析も可能になってきている。農業者へのインセンティブとして、農村生活の困難さを解決するために住民や地域のインフラ管理に展開する農村 DX が始まっている。

現在の農業 DX の課題としては、最先端の本質的な農業データが、農業者や研究者からまだ収集できていない点大きい。データはクラウドにあるものの、農業者個人が利用するためにデータが蓄積されており、心情的な障壁も有りデータ共有できていない実情がある。また、GDPR をベースにした日本の農業データ利活用のガイドライン等は農業者を守る方針になっている事にも原因がある。国や公的研究機関が異業種・異分野から個人に帰属している農業データへアクセスできるように変えていかなければならない。農林水産省の助成制度で得たデータはオープンサイエンスでの利活用できるようにする必要がある。海外ではメジャー企業から提供される農業生産支援パッケージのビジネスモデルが確立している。ドイツの Agri GAIA では、農学研究部分だけでなくデータインフラシステムがサプライチェーン全体の AI 活用を行っている。エコシステム全体の設計を行っている。分野融合研究への応用として農業は環境問題の解決には大きく関わってくるので、農業事業者とビジネスの関係が構築されインセンティブも生まれると考える。

<付録 6 >

令和 4 年 9 月 5 日 第 7 回小委員会話題提供 「DMGMORI における DX 化の取り組み」藤嶋誠（DMG 森精機株式会社） 発表及び議論の概要

生産科学分野では、IoT やセンシング、AI 等の技術革新により開発製品の付加価値を向上させ、高効率な機械の製作や複合加工を可能にする動きが進んでいる。

工作機械メーカーである DMG 森精機株式会社においては、さらにデジタルツインによるテストカット等によって実機加工前にプロセスを最適化することを実現し省エネや自動化による CO2 排出量削減の取組にもつなげている。また、顧客対応においても DX 化を進め、携帯電話網を利用して機械を遠隔管理することで安全に機械の診断・修正・アップデートを可能にするとともに、修理履歴レポートやマニュアル類の一元管理によって顧客の社員の意識改善や稼働率向上につなげる取組を行っている。顧客サポートにおいて得られる様々なデータについては、サービスセンターからのフィードバックや機械の設計に活用するための機械学習の精度を向上させることに有用であるが、顧客の同意を得ながら利用する必要がある。

一方で、インダストリー4.0 の流れの中で工作機械の業界においても標準化が進みつつある。しかしながら、工作機械の CNC のデータを標準化して入出力するフォーマットは整いつつあるが、十分に活用する段階には至っていない。データ共有化に関する議論については RRI において競争分野と共同分野に分けて議論がされている。データの提供先としては、信頼できることが条件であり国又はその代理機関への提供も考えられるが、顧客に対してどこまでフィードバックできるかが重要である。また、集積されたデータから得られた知見については、元のデータに顧客のデータやノウハウが含まれているため、知財として扱う場合は国や公的な機関においてルールが決められていることが望まれる。

（スライドは非公開のため掲載なし）

<付録 7 >

令和 4 年 9 月 9 日 第 8 回小委員会話題提供 「建築分野におけるオープンデータ利用」田辺新一（早稲田大学） 発表及び議論の概要

建築分野におけるデータ等のオープン化については、(1)民間企業との共同研究の場合の契約上のハードル、(2)個人情報保護の観点から難しい等の課題が残っている。オープンデータの成功事例としては、日本建築学会の「市街地風環境予測のための流体数値解析ガイドブックガイドラインと検証用データベース」があり、検証用データとして利用した例や、カリフォルニア大学バークレー校を中心として世界中のオフィスでの在室者の温冷感や室温等のデータを公開するプロジェクトがある。なお、2020 年度から国土交通省が始めた現実の都市をサイバー空間に再現する「Project PLATEAU」や「i-Construction」も進めており、その中で BIM (Building Information Modeling) として、3次元の建築デジタルモデルを設計から維持管理に至るまであらゆる工程で活用する取組を進めている。しかしながら、日本は現場の熟達者が各自の担当範囲について高度な技術で対応する慣習があり、デジタル化・共有化が遅れているという問題がある。近年、現場の熟達者の高齢化が進んでおり、建築物の品質を維持するためにも、BIM 等のデジタル・共有化を進める必要がある。また、公的機関が所有するデータだけでなく、民間企業が所有するデータについても利活用できるように、行政側からのデータのオープン化に関する働きかけが必要である。さらに、研究者にとって、データをオープン化することに対するインセンティブを付与する試みが必要である。

日本建築学会「市街地風環境予測のための流体数値解析ガイドブックガイドラインと検証用データベース」(2007年～)

- ✓ 各種の建物形状を対象としたベンチマークテストに基づいて、ガイドラインが策定され、検証用のデータベースを併せて用意した
- ✓ ガイドラインにおいて利用を推奨している実験データベース、開設以来1万アクセスを超え、国内外の多数の論文でCFD検証用データとして利用



Department of Architecture, WASEDA University
https://www.aij.or.jp/jpn/publish/cfdguide/index_e.htm

5. BIMとは

BIM: Building Information Modeling

- ✓ 3次元の建物デジタルモデルに、コストや仕上げ、管理情報などの属性データを追加した建築物のデータベースを、建築の設計、施工から維持管理までのあらゆる工程で情報活用を行う方法
- ✓ 設計から施工、維持管理に至るまで建築ライフサイクル全体でモデルに蓄積された情報を活用することで、建築ビジネスの業務を効率化し、建築デザインにイノベーションを起こしつつある画期的なワークフロー



我が国では、アナログによる設計・施工が非常に優れているため欧米アジアと比較するとデジタル化が遅れている

AUTODESK RVTより引用 <https://bim-design.com/bim-dx/bim-01-about-bim/>

Department of Architecture, WASEDA University

<付録 8 >

令和 4 年 9 月 29 日 第 9 回小委員会話題提供

「オープンサイエンスを推進するデータ基盤に関する話題提供」

上東貴志（神戸大学）


発表及び議論の概要

社会科学分野では、研究機関によるデジタル資料の公開・利用が進み、主要ジャーナルではデータポリシーの一環として、再現性を確保したデータやコードの公開が進められている。この背景には、社会科学、データ科学、計算科学が融合した分野である計算社会科学の台頭がある。計算社会科学は、twitter 等の Web 上の大規模な社会データを機械学習等の手法で分析し、その結果を解釈する学問であるため、データやコードを公開することで再現性を確保することが慣習となっている。データ公開のニーズが高まる一方で、プライバシーの問題から公開が難しい社会科学系データも少なくない。神戸大学経済経営研究所では、企業の歴史的な原資料を収集・保管しているが、歴史的なデータであっても訴訟リスク回避のために個人情報処理を行う必要があるデータも数多く、そのデータに個人情報が含まれているかどうかを確認し、匿名化処理を行う人材を確保することが課題となっている。

近年、社会科学分野では個人情報を含むデータを取得する際の倫理審査が厳しくなっており、データ蓄積における個人情報の取扱いルールを整備することが必要と考えられている。各分野の個人情報の取扱いルールを比較・照合することで、一般的なルールを策定できる可能性がある。個人情報、権利問題、機密情報等によりデータの公開が困難な分野は、社会科学分野だけではない。公開しにくい分野のデータは過小評価される傾向にあるが、健全な研究を行うためには、すべてのデータを公開する必要はない。オープンではないサイエンスが尊重されることが望まれる。


神戸大学経済経営研究所

- 1919年に神戸高等商業学校商業研究所として創設
- 2019年に創立100周年
- 社会科学系の国立大学附置研究所として最も長い歴史と伝統
- 経済学・経営学の両分野の先端研究
- 鐘紡資料・兼松資料等の歴史的な企業原資料
- 財務諸表、年次報告書、社史等の企業情報の収集・整備
- 新聞記事文庫、鐘紡資料データベース等のデジタル・データ



6

計算社会科学とは？ (Computational Social Science)



社会科学、データサイエンス、計算科学の文理融合領域

20

<付録 9 >

令和 4 年 10 月 18 日 第 10 回小委員会話題提供
「社会科学分野におけるオープンデータの活用：国際関係・人間の安全保障への適用事例を中心に」 阪本拓人（東京大学）
発表及び議論の概要

社会科学分野の国際関係研究においては、データ駆動型研究が普及しつつある。その中心はオープンデータから線形モデル等を生成する研究であり、オープンデータの拡充が普及の背景にあると考えられる。知名度の高いオープンデータとしては、Uppsala Conflict Data Program (UCDP) や Armed Conflict Location & Event Data Project (ACLED) が挙げられる。近年では、個別のサーバーではなく、Harvard Dataverse 等のデータリポジトリを通じたオープンデータの拡充が進んでいる。

阪本先生自身もオープンデータを用いた研究に取り組んでいる。具体的には、国際連合の安全保障理事会における公式討議の公開議事録にワードエンベッティングを施したうえで、ワードクラウドによって可視化することで、主要国の脅威認識の変化や異同の把握を試みている。また、公開議事録はオープンデータではあるものの、スクレイピングや分析に不向きな構造化されていないものであることから、分析の過程で構造化した公開議事録のオープンデータ化を予定している。

社会科学分野においては、オープンデータを明示的に評価するための仕組みが欠落していることから、その確立が課題として挙げられており、オープンデータの紹介を主眼とした論文から構成されるデータジャーナルの創刊が期待されている。

プロジェクトの概要

- ・大規模テキスト分析を用いた各国・各機関代表による安保理の政策討議の体系的分析
- ・データ：1946年から今日までの安保理公式討議（9,000回以上開催）の議事録(proces-verbaux, P.V.)
- ・目的：安保理メンバー（特にP5）の政策志向や認識の定量的な把握、政策志向・認識の一致・不一致が起きる条件の特定
- ・現状：冷戦後(1990-2021)の発言データの構築、単語埋め込み等を用いた主要メンバーの「脅威認識」とその変化・異同の把握

脅威認識の可視化

