

NII Today

National Institute of Informatics News

104
Jan. 2025

P2 「データの時代」を生き抜く
江村 克己／山地 一禎

P8 研究活動の変革を目指す
松原 茂樹／甲斐 尚人／長岡 千香子／平木 俊幸

P12 研究データ基盤活用の先駆を目指して
小野 寛太／木村 映善／徳地 直子／菊池 信彦

P16 ポリシーは起爆剤、マネジメントは推進力
長井 圭治

Essay データ利活用研究コミュニティの創成に向けて
田浦 健次朗

[特集]

研究データエコシステム

分野を超えた研究データの連携と循環



特集

研究データエコシステム

分野を超えた研究データの連携と循環

「データの時代」を 生き抜く

研究データエコシステム構築への道

江村 克己

EMURA, Katsumi

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推進委員会 委員長
福島国際研究教育機構 理事

山地 一禎

YAMAJI, Kazutsuna

国立情報学研究所
コンテンツ科学研究系 教授
同オープンサイエンス基盤研究センター長

データ駆動型研究や生成AI(人工知能)など大量のデータを活用してイノベーション創出を目指す研究が世界の潮流となっている。また論文だけではなく研究データをインターネットを介して公開して広く共有する「オープンサイエンス」の考え方も国際的に広がっている。「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」は、分野を問わず日本全国の研究者がデータ管理・公開・利活用に取り組める共通の基盤を提供するのが目的だ。本事業推進委員会の江村克己委員長と事業の中核である研究データエコシステム構築事業推進センターの山地一禎マネジメントチームリーダーに進捗状況などを聞いた。

聞き手

滝 順一

TAKI, Junichi

科学ジャーナリスト
(元日本経済新聞論説委員)

大学間での認識共有と コミュニティの拡張

— 2022年度の「研究データエコシステム構築事業」発足からおよそ2年。2026年度末のゴールまでの折り返し点に差し掛かっています。これまで2年間の評価は、いかがですか。

山地 国立情報学研究所(NII)は研究データの利活用促進を目指しNII RDC(リサーチデータクラウド)の構築に取り組んできました。NII RDCはデータの管理・公開・検索の3つの基盤で構成された研究データの包括的な活用基盤です。研究データエコシステム構築事業では、NII RDCの機能を「活用」「信頼」「蓄積」など7つの側面からさらに機能拡張、高度化する研究開発にNIIとして取り組んでいます(図1)。同時に、理化学研究所、東京大学、名古屋大学、大阪大学の4機関が、データ活用・公開のルール・ガイドラインの整備や人材育成など、研究機関が研究データ基盤を活用していく上で必要な環境、いわばエコシステムの整

備を推進しています。

大学間の連携、 地域コンソーシアムの結成

私の実感として、2年経って基盤づくりにより強いリアリティが求められだしました。たとえば、大学として

「こういうデータポリシーを持ちたい、そのためにはこれがなされていないと組織としてのデータ管理が実現できないので、まずはこれをやりたい」とか、「人材育成もしなければならない」という意見です。研究データの管理・活用で、大学側としてのニー

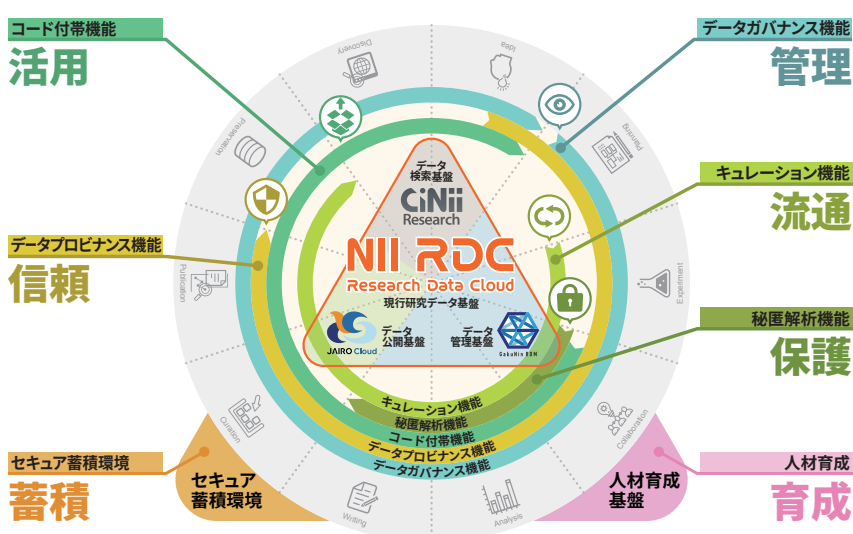


図1 NII RDC(Research Data Cloud)の機能拡張概念図
NIIは、研究データエコシステム構築事業における中核機関群のひとつとして、また、研究データ基盤高度化チームのリーダー機関として、研究データ基盤の機能実装を目指し、NII RDCを7つの側面から機能拡張する役割を担う。

ズが徐々に増えてきた状況にあります。その具体化に向けて、大学間の連携や地域コンソーシアムの結成が進んでいます。名古屋大学を中心に東海地区で2023年度にコンソーシアムが誕生したのを皮切りに、北陸、中国四国、九州、北海道、東北と各地に波及し、地域コンソーシアムがいくつも生まれようとしています。これは事業を推進する側から一方的に仕掛けたことではありません。同じような問題意識を抱えた大学から生まれた動きでもあります。

NII RDCの機能高度化はこうした現場のニーズを先取りする形で進め、大学側のニーズの高まりに応じ事業をアジャストし、ぴったりピースが組み合わさるようになりつつあります。

国内研究者が活用できる 共通のデータ基盤が必要

—なぜ今、研究データエコシステム構築事業が求められるのでしょうか。事業発足の背景について改めてご説明ください。

江村 これまでも「データ駆動科学

の時代」などと指摘されてきましたが、今、データの時代が本格的に到来しています。とりわけ近年の生成AIの登場と急速な普及でデータの価値がこれまで以上に高まってきています。

また生成AI登場以前から、オープンサイエンス、オープンデータの時代だとすでに言われてきました。論文をきちんと出せばよかった時代から、論文を支えるデータもちゃんとオープンにしていこうという形に科学のコミュニティの構造が変わってきています。

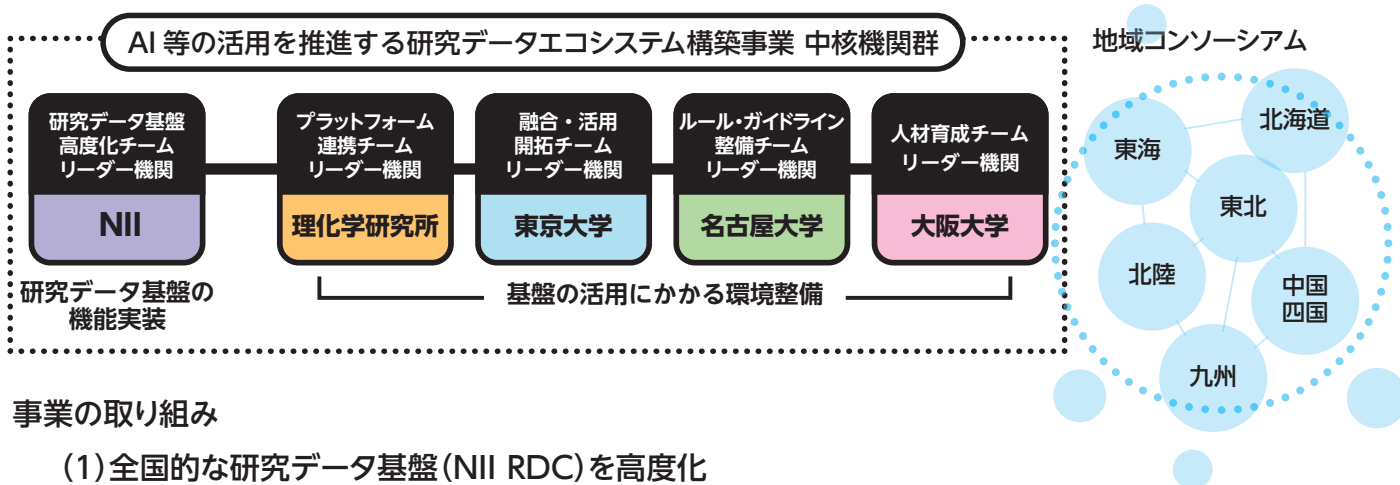
2024年7月にイタリア・ボローニャで開かれたG7(主要7カ国)科学技術大臣会合はオープンサイエンスの拡大をうたった共同声明を発表しました。文部科学省は研究成果のより幅広い活用を狙って、学術論文の即時オープンアクセスを進める方針を掲げています。多くの研究者が、オープンデータの重要性を認識し、データをうまく活用していくよう求められています。

そんな時代にあって、研究データの活用という観点では、率直に言っ

て日本は世界の中で必ずしも進んでいるとは言えない状況です。データ活用ができる基盤をしっかりと作っていかなければいけないという社会的・学術的な強い要請がある状況下で、ともすると大学ごとにバラバラにいろんなものを作ってしまうようなことにならないよう、分野を問わず日本全国の研究者が活用できる共通の研究データ基盤をNIIをコアとして構築することが、まず一つ目の大きな背景です。

異なる領域を結ぶ データ連携プラットフォームが 研究データエコシステム

もう一つはデータ連携によって新しい価値を生み出していきたいということです。これまでは研究者が自分の研究領域の中でデータを活用するという状況が主流だったのですが、異なる領域が結ばれるデータ連携によって新しい価値が生まれてきます。それを可能にするようなプラットフォームが必要になります。それを「研究データエコシステム」と称しているわけです。これはこの事業に



事業の取り組み

- (1) 全国的な研究データ基盤(NII RDC)を高度化
- (2) 研究データ基盤の活用を促進するための環境整備
- (3) オープンアクセスの推進に向けた機能強化など



江村 克己

東京大学大学院修了。光通信の研究者として日本電気(NEC)入社、同社CTO(チーフテクノロジーオフィサー)、シニアアドバイザーなどを経て現職。総務省情報通信審議会委員、科学技術振興機構(JST)AIPネットワークラボ長、BIAC (Business at OECD)イノベーション・技術委員会副委員長など多数を歴任。1987-1988米国Bellcore客員研究員。工学博士(東大)。

とってはいわば究極のゴールとも言えるもので、そこに向けて近づいていこうと考えています。

—データ活用を通じた新しい価値の創造は日本の研究力向上にも重要ですね。

江村 ものすごく大事なことだと考えています。日本から発表されるトップ1%やトップ10%論文が減って日本の研究力が落ちていると指摘されているのですが、私は学際的な研究が弱いところに本質があるのではないかと考えています。文部科学省の科学技術・学術政策研究所(NISTEP)の分析によると、世界で注目される学問領域の数が増えています。2002年の分析では、国際的に注目を集める研究領域の数は約600でしたが、2020年の分析では900領域を超えています。増えているのは物理学とか生物学とか伝統的な学問領域を超えた、いわゆる学際的・分野融合的な領域です。

米国や中国はそうした新しい領域でどんどん論文を出し研究が進歩していますが、日本はそこが弱いようです。日本は伝統的な領域で強みを発揮していますが、学際的・分野融合的領域での注目度の高い論文

の数が伸び悩む傾向にあります。全体として日本のトップ論文が減る根っこはそこにあるのではないかと思います。

つまり日本は新しい学問領域を作りきれてない。その点でこのプラットフォームは非常に重要です。もちろんプラットフォームがあれば学際的領域が生まれるかと言われれば、そんなに単純なことでもありません。ただプラットフォームでいろんな人がそこで混じり合いつながるようになってくるのは間違いありません。

ユースケースによる コミュニティ形成が出发点

—研究領域が異なれば、対象データの性質も量も違います。データポ

リシーづくり一つとっても、なかなか学内の合意が得られないという課題も耳にします。人材育成が大きな課題ではないでしょうか。

山地 データ連携による学際的研究領域の創出や新しい価値の創造は、江村委員長が指摘されたように事業の最終的な目的と言えると思います。その前段としてデータの活用を目指したプラットフォームを構築してきたところです。次に求められるのは、これを使いこなす人材の育成です。大学でデータを管理し活用するためのツールを研究者に紹介するコーディネーター的な役割を担う人たちも必要とされるでしょう。海外ではそうした役割を専門とする人がいて学際的な研究を促した



りデータ駆動型の研究を支援したりしています。

事業の一環でユースケースを公募し33の取り組みがスタートしています(2025年1月現在)。同じ基盤を使う、言い換えれば「同じ釜の飯を食う」コミュニティが生まれ、それらを支援するうちに学際的な共同研究が進むといった成果もこれから出てくると期待しています。

江村 「課題」ではなく、進むべき「道」が見えてきたと捉えるべきでしょう。多様な研究者が同じプラットフォームで活動し始めるのがスタートポイントです。まずはコミュニティづくりです。コミュニティの人たちがこれとこれをつないだら価値が生まれるのではないかと提案できる段階になってくると本当に世界観が変わってくると思います。

データ連携で 学際的研究領域が生まれる

江村 推進委員会は事業の成果を評価することも役割ですが、推進の立場から「伴走する」とも発言してきました。委員の構成は多様で、法学など文系の委員もいらっしゃる。「私たちにも使えるようにしてください」という声が上がります。研究データエコシステムの価値を多くの人に理解してもらわないといけません。データ連携で学際的な研究領域が生まれる、そこで新しい価値が生まれるのだということの意味、その重要性を伝えることがこの事業にとつてとても大事だと考えています。

山地 補足しますと、オープンサイエンスにしてもデータガバナンスにしても研究機関は「義務的なもの」と捉えがちです。研究者にとつ

て「やらされる仕事」という側面があります。しかしこの事業のユースケース創出公募で手を挙げて取り組んでいる人たちは義務からではなく研究促進のため取り入れようとしています。

米国では全米科学財団(NSF)や国立衛生研究所(NIH)あるいは個々の大学がデータ基盤作りを進める一方、連邦レベルでリサーチ・インフラストラクチャーにもっと投資をしなければならないという議論があります。欧州連合(EU)は欧州オープンサイエンスクラウド(EOSC)があって巨額の投資をしています。ただNII RDCのように国内を全てカバーできるような一体化したものではありません。日本が先駆けている面もあります。これを義務のためだけではなく、研究促進のためのサービスへと発展させ、世界をけん引し続けるのが我々の役目です。

——事業には終わりがありますが、システムは永続します。10年、20年先を見越して推進委員会としてなすべきことはありますか。

江村 長期の視点では、研究データエコシステムのサステナビリティが気になります。どういう形でこのシステムが自立していくかです。これはNIIだけで解ける問題ではありません。参画する大学や、将来的には企業にも入ってもらった方がいいと思いますが、受益者がシステムを回していく負担をしていくことになると考えます。システムによって新たな価値が生み出されるのであれば、皆が参入してくるはずですから、そこをどうデザインするかが大事です。あと2年ほどでデザインしきれるか

と問われれば、正直なところ容易ではないとは思いますが、どういう形で持続的なものにするか、推進委員会としても提言のようなものを出していかなければと考えています。

聞き手からの ひとこと

日本政府が「信頼性のある自由なデータ流通(DFFT)」という旗印を掲げたのは2019年。その後の新型コロナ感染症への対処を振り返ると「道遠し」の感があった。研究機関でのデータ管理や利用はもはや待ったなしの状況とも言える。認識を共有するコミュニティの誕生は心強い。これをいかに広げ深めて持続させるかが課題だ。



滝 順一

科学ジャーナリスト

早稲田大学政治経済学部卒業後、日本経済新聞社に入社。1980年代半ばから科学技術の研究開発現場と科学技術政策の立案プロセス、気候変動、エネルギー政策などを取材してきた。日本経済新聞社退社後、日本科学技術ジャーナリスト会議事務局長を務める。

研究データエコシステムの構築は、どこまで進んでいるのか

国立情報学研究所
研究データエコシステム構築事業推進センター
プログラムマネジャー 中野 恵一

「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」は、国立情報学研究所（NII）が4つの研究機関・大学とともに取り組んでいるプロジェクトです。期間は2022年～2026年、2025年1月現在、中間点を少し過ぎたところです。目的は、今すでにある研究データ基盤（NII RDC）を充実させつつ、研究データの利活用を促進することによりイノベーション創出に貢献できるエコシステムへと進化させることです。

最大の意義は、データの共有と利活用が促進されることで、学術研究の透明性と再現性が飛躍的に向上する点にあります。AIや機械学習を活用したデータ解析が高度化し、データ駆動型の科学研究が進展している現在、適切なデータ管理と利活用は必要不可欠です。データは誰でも簡単に見つけられ（Findable）、必要に応じてアクセスでき（Accessible）、他のシステムやデータと連携でき（Interoperable）、再利用しやすい形で提供される（Reusable）ことが重要です。さらに、秘密情報が確実に保護されることも欠かせません。そのような世界の実現に向けて本事業では、「NII RDCの高度化」に加えて、「プラットフォーム連携」「ユースケース創出」「ルール・ガイドライン整備」「人材育成」に取り組んでいます。

「プラットフォーム連携」では、特にオープンサイエンスが進んでいる分野に見られるような、その分野のための基盤システムや、さまざまな機関が有している既存システムとNII RDCとを、有機的に連携させるために必要な技術的あるいは運用面での工夫が検討されています。

「ユースケース創出」では多様な研究分野のテーマが累計で33件（2025年1月現在）採択され、データ活用の実践例として蓄積し始めて

います。これらは、さらに新しい研究アプローチの発見につながり、その研究の発展が次に続く研究を充実・加速させる生態系が実現するでしょう。

「ルール・ガイドライン整備」には、多くの機関が高い関心を向けています。第6期科学技術・イノベーション基本計画では、機関リポジトリを持つ全ての国立大学・国立研究開発法人などに対し、2025年までにデータポリシーを策定するよう求めており、締め切りが迫っているからです。それぞれの機関の事情に合っていないと役に立ちませんが、ゼロから作り上げるのは大変です。それでも、先行している機関がそれぞれ自機関の事情に合わせたデータポリシーを公開しており、策定の取り組みの中で得たノウハウも蓄積・共有し始めています。研究データ管理を全国的に展開するための素地は、整いつつあります。

「人材育成」の取り組みでは、研究データ管理に必要なスキルや知識を効果的に習得するためのプログラムを開発しています。研究者とともに研究データ管理を担う人が育成されれば、データリテラシーが向上し、エコシステムの活用も促進されるでしょう。

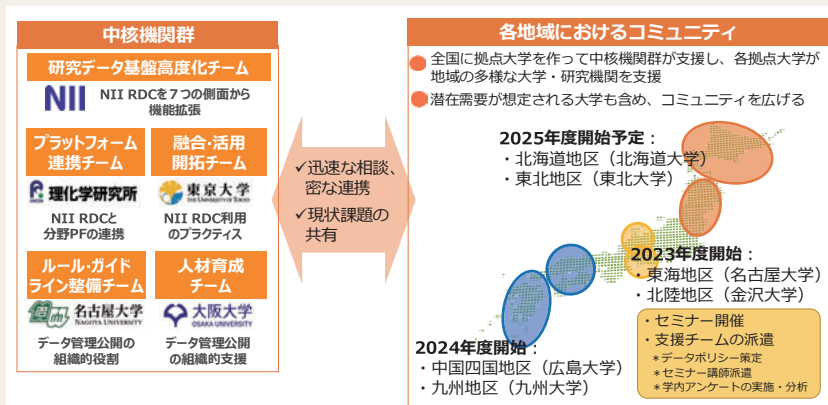
さらに2023年度からは、本事業の中で「研究データ管理スタートアップ

支援事業」を開始し、地域の研究機関や大学が研究データ管理を進めるための初期支援を行っています。すでに地域コミュニティの力とリソースを活用した独自のユースケースが生まれ、それぞれの地域のニーズに応じたデータ管理の実践が始まっています。今後はさらに多くの地域コミュニティが立ち上がり、全国的なネットワークとなることで、いずれは全国の研究機関が連携して進化し、持続可能な研究データ管理基盤を形作ることでしょう。

研究データエコシステムは、研究者だけのものではなく、一般社会にも大きな価値をもたらします。このエコシステムの構築によって、地域課題の解決や産業の発展、災害対応の迅速化、さらには医療や福祉の向上など、多様な分野での具体的な成果が期待されます。研究データの利活用が進むことで、社会全体がその恩恵を享受し、より豊かで持続可能な未来の実現に寄与することを目指しています。

今後もさまざまな方々と共に歩み、科学研究の発展に貢献できるよう、研究データエコシステムの構築という困難で重要な目標に向かい続けたいと思います。

研究データ管理スタートアップ支援事業の体制と展開状況



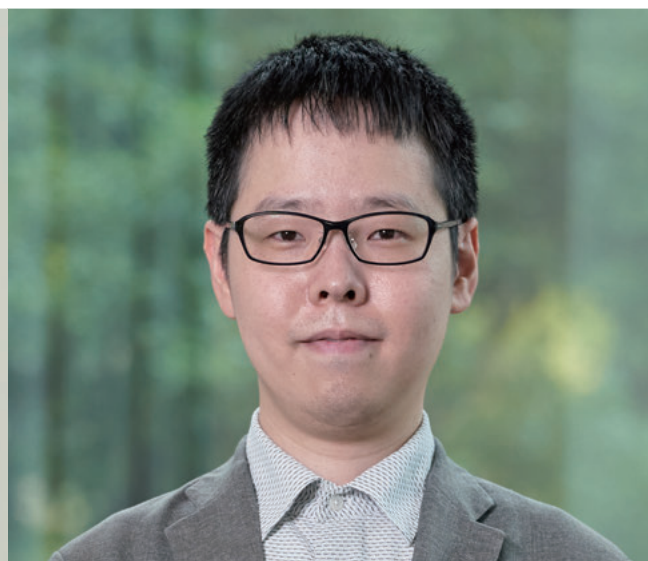


研究活動の 変革を目指す

ルール・ガイドライン整備チーム
人材育成チーム

松原 茂樹／甲斐 尚人
長岡 千香子／平木 俊幸

文部科学省が取り組む「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」は、デジタル技術とデータ活用による研究活動の変革—研究デジタルトランスフォーメーション(研究DX)を推進する事業で、「ユースケースの形成・普及」「データ共有・利活用の促進」「研究デジタルインフラ等の効果的活用」を一体的に推進する。事業は、国立情報学研究所(NII)と共同中核機関の理化学研究所、東京大学、名古屋大学、大阪大学によって2022年に開始した。その中で今回「データ共有・利活用の促進」のため、適切な研究データ管理支援、データマネジメント人材育成支援を目指す、「ルール・ガイドライン整備チーム」および「人材育成チーム」に、現在までの進捗、課題、展望について聞いた。



■「研究データ管理」の普遍化を目指して

——研究データエコシステム構築事業での活動の概要を教えてください。

松原 研究データエコシステム構築事業ではNIIのほかに4つの共同実施機関が参画しています。名古屋大学は「ルール・ガイドライン整備チーム」のリーダー機関として、研究データの取り扱いに関するルールを国内の大学で整備、普及させ、全国的に統一的なフレームワークの下で研究データが扱われるようになることを目指しています。

大学における研究データの取り扱いは、現状では、研究グループや研究者の裁量で決まることが多いと思います。研究者にとっては自由度が大きく、一見、望ましいように見えますが、行動するたびに扱い方を判断しなければならないのは非効率です。

ルールやガイドラインがあらかじめ定められていれば、研究者が研究データの取り扱いで迷う必要がなくなり手続きが効率化されますし、グループで研究データを共有する場合にもデータの意味や管理方法などのあいまいさも解消されます。また、研究支援者にとっても支援方法を明確化できるなど多くのメリットがあります。

大学にとっては、グループや研究者ごとに異なっていたデータの取り扱いや置き場などを統一することで、研究データのガバナンス、つまり、どのようなデータがどこに存在するかを把握しやすくなります。データ利活用の施策も打ち出しやすくなり、大学にとってもメリットが大きいと思います。さまざまな分野のデータがひとところに集積するようになると、デー

タの利活用を通じて人と人、研究者同士の結び付きが強まり、研究活動の厚みが増すのではないのでしょうか。データの価値とはそういうところで高まっていくのだと思います。

平木 私は「ルール・ガイドライン整備チーム」と協力し、データガバナンスを支えるための機能の開発を行っています。研究データの管理を適切に行う上で、現在のデータ管理の状態を把握することが重要です。データガバナンス機能は、1.データ管理上何を守らなくてはいけないか(ポリシー・ルール)を見える化、2.今のデータ管理状態を見える化、そして3.今のデータ管理状態がポリシーやルールに沿っているかどうかの検証の支援——という大きくこれら3つの柱の実現を目指しています。

甲斐 まずは研究データ管理の重要性を広くしっかり浸透させていくことが重要です。そこで「人材育成チーム」では、研究データ管理を正しく理解した研究者や研究支援者を育成する環境の構築を目指しています。具体的には、まずデータ管理に関する教材を開発し、学習管理システム(LMS)を通して全国の研究機関に水平展開することを目指しています。そして職種別学習カリキュラムの開発、さらにはラーニング・アナリティクス(LA)基盤の構築による個々人に適した効果的な学習環境を通して、研究データエコシステムのサイクルに適切にトルクをかけられる人材の育成を目指しています。

長岡 大阪大学が作成し、オープンアクセスリポジトリ推進協会(JPCOAR)で標準化した教材を全国に展開する際のプラットフォームとしては、NIIが提供している学習管理システム、

「学認LMS」を利用しています。学認LMSでは、学習コンテンツの提供、改善のための学習分析機能の提供、関連するサービスとして教材再利用のためのリポジトリ提供など、これまでに構築してきたプラットフォーム機能があり、これらを通して人材育成に関する支援を行っていく体制です。

研究データ管理(RDM)の学習は、研究データを適切に扱い、再利用できる形にするために必要な知識やスキルを知ることができますので、一部の分野の研究者だけのものではなく、研究データを取り扱う全ての研究者にとって、データを取り扱うための必須の資格という位置付けにしたいと考えています。そのために教材の作成および普及を進めていращやる甲斐先生ほか大阪大学の取り組みをプラットフォームでサポートするのが私たちの役割です。

■事業推進の現状と課題

——現在の活動の進捗、課題などについてお聞かせください。

松原 大学における研究データ管理の推進は、多くの大学が相互に参照し合いながら進めていくのがよいと考えます。実のところ、「研究データエコシステム構築事業」が開始された2022年夏の時点では、ルール・ガイドラインの出発点とも言える「研究データポリシー」を策定している大学は10校に満たない状況でした。国の施策として「2025年までに大学がポリシーを策定する」という目標が掲げられていますが、2024年末の時点では90校程度に達しており、着実に進んでいる印象があります。

もう一つ、研究データの公開に関しても、全国の大学で機関リポジトリの整備が進められています。従来、



松原 茂樹

MATSUBARA, Shigeki
名古屋大学
情報基盤センター 教授

論文の公開が主だったのですが、研究データに関しても、リポジトリで公開するためのメタデータの基準づくりが進んでいます。

一方で実現が難しいと感じるのは、すでに研究グループ内で独自のスタイルが浸透している場合です。それらはさまざまに最適化された結果として定まったものでしょうから「現在の方法をやめて、今後は大学のガイドラインに従う」というのは簡単ではありません。研究データの公開についても大きな負担だと感じる研究者や研究支援者は多いと思います。それもまた難しさの一つです。

NII研究データ基盤の発展や普及が、この問題の解消に大きく寄与するものと期待しています。

平木 データガバナンス機能に関しては、特にここ2、3年はコア部分の機能、つまり「ポリシーをうまくシステム側で判定できるように設定する」、および「設定されたポリシーに則って、データ管理の状態を検証する」という機能の開発を進めています。

しかし、実際のルール・ガイドラインを、我々のデータガバナンス機能側でポリシーとして適用することに非常に難しさを感じています。現状、ルール・ガイドラインとして整備されているものは、人間には理解できても、システムからすると『読みづらい』、つまり機械可読性が低いからです。「それらの機械可読性を向上するにはどうすればいいだろうか」

「どのようにロジックを組めばいいのだろう」と悩むこともあります。このあたりは松原先生と相談しながら、詰めていく必要があるだろうと思っています。

長岡 人材育成に関して、教材の作成は比較的順調に進んでいるように感じています。反面、今後について、やはり「どのように各機関に研究データ管理の文化や教育を浸透させていくのか」という点が最も大きな課題だと感じています。

今後も研究データ管理に関する知識や関連する教材は増えていくと思いますが、たとえばそれらの活用を義務化するとしても、では、どのレベルまでの学習を受講者に求めるのか。また、義務化と言っても、学習者である研究者や研究支援者が「嫌々ながらやる」ものにはなってほしくないわけで、それらを学び使うことで『いいこと』がある、と感ぜられるように持っていきたい。義務化の一方で、そのような、学習者が能動的に楽しく学ぶことができる路線も模索していきたいと考えています。

甲斐 教材の作成と展開に関しては、当初想定していたよりも前倒し

で進めることができていると思います。これは、我々大阪大学とNIIとが、チームとして、逐一情報を共有できているというところが大きいと感じます。

特に、教材の展開について、NIIの学認LMSの活用をまず前提に置いたことは、展開先の負担軽減につながっていると思います。

長岡 NIIとしては、この「研究データエコシステム構築事業」が始まる以前から、各機関の教育コンテンツの開発・維持に関するコストを軽減すること、また、一部の教育者・研究者が作成した良質な教育コンテンツを全国のレベルで共有できるようにすることを目的として、教育コンテンツを共有・学習できるプラットフォームである学認LMSを構築してきました。いわば“種”をずっと育ててきたわけですが、そのプラットフォームが、今回うまくマッチしたと言えると思います。

■全国展開・水平展開のビジョン
——そのような課題を踏まえ、今後の展望についてはいかがですか。

松原 ルール・ガイドライン整備チームとして、まずは2024年度中に研究データのガイドライン整備を行い、2025年度からそれを全国に展開していく予定です。それと併せ、先ほど平木さんからもあったように、整備したガイドラインを、NII研究データ基盤に機能としてどのように盛り込ん

平木 俊幸

HIRAKI, Toshiyuki
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



甲斐 尚人

KAI, Naoto
大阪大学
D3センター 准教授



でいくか。これをNIIの皆さんと協力しながら進めていければと思います。

平木 策定したルールやガイドライン、ポリシーを、研究データ管理に活かすための仕組みの開発を進めていると考えています。我々が今作っているデータガバナンス機能の中で、系統的にそのポリシーが機能しているか。また、そのポリシーの下で研究者の管理が行えているか。これらをきちんと検証できる形にしていけることが重要だと思っています。それらの実現のためにも、ポリシーなどの機械可読性の向上の検討も進めていきたいと考えております。

甲斐 人材育成に関しては、先ほど申し上げた通り、我々が現在開発している教材などの育成の仕組みを全国展開・水平展開していくことが目標です。実際には、研究データのエコシステムをうまく回すための人材育成は、分野や領域によっても違いがあるのですが、研究機関がまず理念浸透という第一歩を踏み出すという点では、おおよそすべきことは共通していると思います。

今後、なるべく効率的に大きな成果を上げるためには、全国の研究機関・大学が共通で導入できる人材育成環境の構築が求められると思っています。また、他チームとの関係では、

現在、名古屋大学との連携も進めています。

昨今、研究者には、このデータ管理のみならず、研究倫理や情報セキュリティなど、さまざまな教育が課されています。そうした中で、本来の研究により多くの時間を確保してもらうためにも、できるだけ負担なく学習していただくというのが必須になるかと思っています。負担の少ない学習を実現するためには、たとえば今回開発を進める教材を誰もがしっかりと全て見る必要はないので、職種など、それぞれの必要性に応じ個別最適化したカリキュラムを作り、必要な部分だけを学んでもらうといった工夫も重要だろうと思います。これは、研究者や研究支援者それぞれが深く関与する研究データライフサイクルの各フェーズにおいて求められるルー

ルやガイドラインの落とし込みによって実現を目指したいと思っています。こういった点で名古屋大学と連携を図ろうとしています。

長岡 人材育成の展望としては、現在、いろいろな大学で導入が始まっているオープンバッジを活用し、どんな知識やスキルを持つ人材がどこにいるかを可視化する仕組みも視野に入れていきたいと思っています。

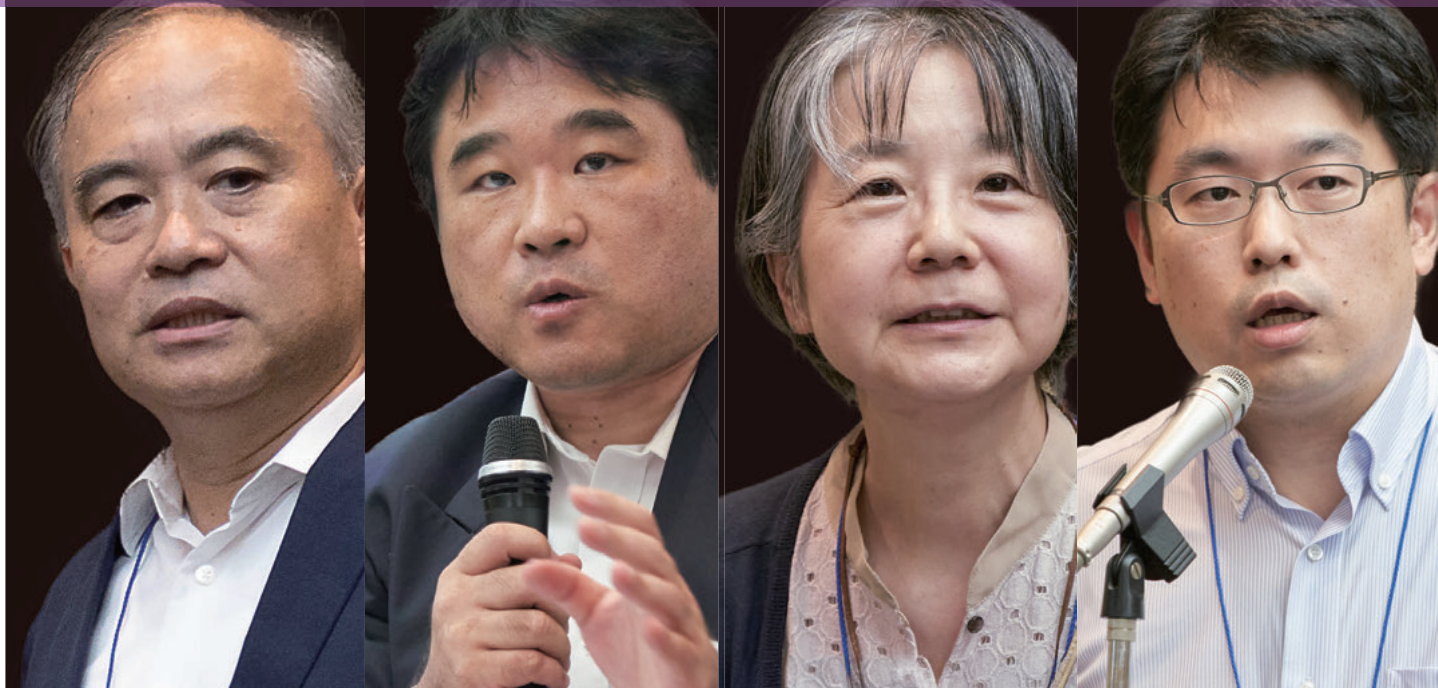
あとは、甲斐先生の仰っていた「個別最適化」の推進ですね。各チーム各機関ごとに、置かれた状況はさまざまだと思います。我々としては、できる限り統一化された手を付けやすい教材をパッケージとして提供することを考えていますが、最初に手を付ける教材としてはいいとしても、それとは別に、「各機関が、それぞれ自分たち用にカスタマイズできる」仕組みづくりも進めていきたいと考えています。

松原 「ルール・ガイドラインの整備」と「人材の育成」は、大学における研究データ管理の両輪です。ルールを作っただけで終わるのではなく、適切に浸透、運用させていく人材づくり、そしてそのための教材も不可欠です。両チームの関係性を強化し、今後も活動を推進していきます。

長岡 千香子

NAGAOKA, Chikako
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





研究データ基盤活用の先駆を目指して

小野 寛太

ONO, Kanta
大阪大学
大学院工学研究科 教授

木村 映善

KIMURA, Eizen
愛媛大学
医学系研究科 教授

徳地 直子

TOKUCHI, Naoko
京都大学 フィールド科学
教育研究センター
森林生態系部門 教授

菊池 信彦

KIKUCHI, Nobuhiko
国文学研究資料館研究部准教授
同古典籍データ駆動研究センター
副センター長

「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」では、全国的な研究データ基盤の活用やデータ連携を進めていく上で、その先駆となるべきユースケース創出に向けての課題提案を公募している。

一言でデータ基盤の活用と言っても、抱える問題意識、活用に向けてのアプローチは、分野により、研究のテーマにより、さまざまな違いがある。ここでは、「ユースケース創出事業(以後、ユースケース事業)」採択プロジェクトの中から、それぞれ異なる分野の4人の研究者に、事業応募のきっかけ、進捗、課題、抱負と今後の研究データのあり方、この事業を通しての、国立情報学研究所(NII)への提言を聞く。

研究データ基盤に関する 問題意識

——今回の「ユースケース事業」ご参画の経緯、研究概要をお願いします。

小野 我々の提案は、カテゴリーとしては「実験」になります。我々は材料科学の分野で、大きく分けて2つの研究を行っています。一つは、大型の施設を使い、放射光や中性子に関し大

規模に計測データを取得、解析し、そこから何か新しいものを見つけようというもの。もう一つは「自動自律実験」というものです。これは人の手を介さず、ロボットに実験を任せるのですが、こちらも連続的に大量のデータが出てきます。

しかし現時点では、入り口に過ぎない「大量のデータを取得する」方ばかり目が向いて、本来重要である

はずの、それら大量のデータを素早く解析し、次の実験に活かしていくというループが、いまひとつしっくり行っていない。それが前から気になっていました。

そんな折、NIIでちょうどデータ基盤に関する取り組みが始まるということを知り、応募した次第です。応募にあたっては、データを研究者のコミュニティで共有し、新たな知識の創出

に結び付けるためのプラットフォームの構築というところまでも含めて提案させてもらいました。(図1)

木村 私の課題提案は、「国際的なRWD(リアルワールドデータ)※1研究を実現する医療情報分析基盤の検討」をテーマに掲げています。(図2)医療リアルワールドデータとは、日々の医療の実臨床で得られるさまざまな医療データの総称です。

私自身は医療情報学が専門で、大学で教授を務める一方、学部付属病院では医療情報部長という立場で電子カルテシステムの運営管理に携わっています。そんな中、多くの先生から「電子カルテからデータを取り出して使いたいね」というリクエストが寄せられるのですが、実はこれが一筋縄では行きません。取り出してすぐに使えるものではなく、個人情報秘匿加工のプロセスを踏むなど、そのマネジメントが重要になります。

こうした管理と活用をスムーズに行えるよう、かねてデータ環境の基盤整備に関して取り組んでいたのですが、どうも研究ではなく、授業とか業務とかの範疇であろうと判断されてか、なかなか公に取り上げてもらえませんでした。そのような状況で、今回この事業に応募し、プロジェクトが採択され進めているところです。

徳地 私が属している京都大学フィ

ールド科学教育研究センターは、20年ほど前、大学の改組の時に「森里海連環学」という新たに設定された学問領域をミッションとして生まれた組織です。これはその名の通り、森から海までを一つのつながりとして捉えて考え、健全な国土を保全していこうというものです。

私自身はもともと森林生態学が専門なのですが、以来、私も含めてセンターの研究者は、ももとの自分の分野よりも視野を広げて研究をするようになっていきます。今で言う超学際研究というものを、その頃からやっているわけです。

ただ、我々自然科学系の研究者は一応定量的な数値データで物事を考えることが多いのに対し、社会科学系はもっとナラティブなアプローチになることもある。そうした異なるデータの“作法”を、どうやって一元的に捉えていくか。それが常々悩みだったのです。そんな話を、情報学の先生にお話ししたところ、こちらのプロジェクトを紹介していただいた。それが応募のきっかけです。

たまたまその時、ご縁があって岐阜県飛騨市の森林をテーマに研究を行っていたこともあり、「飛騨市においてデータを基に、これからの飛騨市を考えることができるようなプラットフォームを作りませんか?」といった課題

を設定させてもらいました。(図3)その先について言えば「異なる分野の人が一緒に考えられるプラットフォームを作る」ことが究極の目標です。まあ、現時点ではまったくそこまで行っていないのですが。

菊池 私が所属する国文学研究資料館(国文研)では、「データ駆動による課題解決型人文学の創生～データ基盤の構築・活用による次世代型人文学研究の開拓～」という、文部科学省の大規模学術フロンティア促進事業のプロジェクトを2024年度から2033年度までの10年計画でスタートさせたところなのです。この中で、人文系の研究データを管理し、それをオープンデータとして公開、流通させていくことは喫緊の課題となっています。「ユースケース創出事業」については上司の紹介で知ったのですが、人文系の研究データの管理・公開に関しての経験を積むにもいい機会であると考えています。

課題提案のタイトルは「古典籍テキストデータを活用したデータ駆動型人文学のための研究資源構築プロジェクト」。過去10年にわたって構築を進めてきた国書データベースを利用し、参加メンバーそれぞれの研究関心の元で、特色のある古典籍のTEI/XMLデータ化に取り組んでいます。そこで試行錯誤して得られた知見や

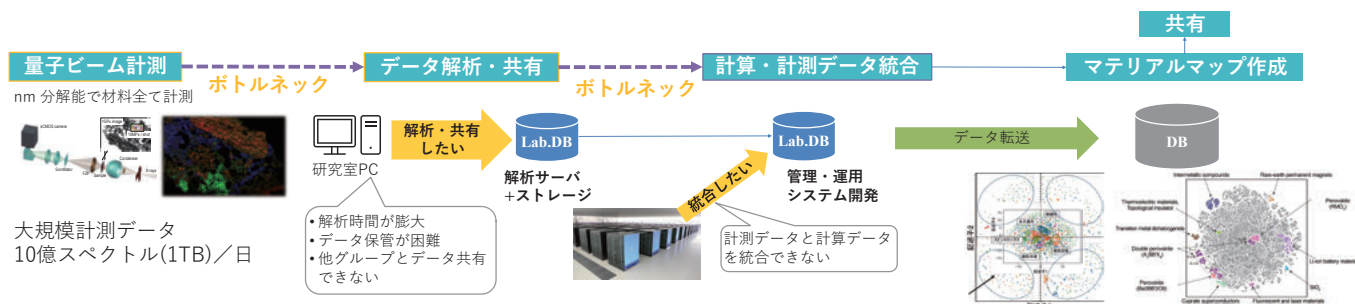


図1 「計測・解析・共有・公開を通じた知の創出:NII RDC、GakuNin RDM、mdxを活用することにより、大規模計測データの計測・解析・共有・公開を通じた知の創出のためのエコシステム構築を目指す」研究データエコシステム構築事業シンポジウム2024での小野教授ポスター発表表。

データを、NII RDCを使い共有、公開していくことで、これからの古典籍の研究資源化を目指しています。(図4)

利活用のレベルも さまざまな NII RDC

——「ユースケース事業」の条件、NII RDCの活用について、活用状況、メリット、課題などお聞かせください。

木村 NII RDCの中でも、現在、私の研究計画では、特にその管理基盤であるGakuNin RDMを利用しています。GakuNin RDMに関しては、特に研究の公正性とトレーサビリティの担保という点でありがたみを感じています。

研究においては過去プロセスが透明であるかどうかは非常に重要なのですが、GakuNin RDMにはタイムスタンプの機能があり、信頼性の高い認証プラットフォームと連携をしています。第三者的に信用を担保できるアーキテクチャとして、非常に有用な仕組みではないかと思っています。一方で、患者さんの立場からすると、そのような多くの人々が利用する基盤に、個人情報である医療データ、研究データを置いて大丈夫なのかと、抵抗感を覚えることもあるかと思いま

す。これについては突き詰めれば信頼関係の問題ではありますが、安心感を高めるための啓発活動のようなものも必要かと思います。これが、私が思い付く課題の一つでしょうか。

菊池 我々の研究の中では、NII RDCは共同研究メンバーのTEI/XMLマークアップデータの保存と共有、そしてそのマークアップ方針の集約場所として利用しています。現時点では扱うデータが小規模なものである場合が多く、共有も常時行う必要がなく研究会の際に共有できればいいために、商用ストレージやメール共有でも十分になっています。その点では、ややNII RDCの恩恵を感じにくいのが正直なところですが、NII RDCを利用するメリットは間違いなく「安心感」にあると思います。従来の商業クラウドサービスは永続的に存在する保証がない中で、NIIによるシステム提供は、「安心して長期に利用できる」と思わせてくれます。

一方で、不満点や改善点は、まさにそれと裏表の関係にあります。究極の安心感は存在や利用を意識しないことにありますが、NII RDCの利用に際しては、これまでの研究習慣になかったことを行い、慣れない操作をし

てデータを登録し、そのバージョンを管理せねばなりません。これまで商用ストレージサーバでボタン一つでできていたデータ保存と共有が、NII RDCとなると、メタデータの管理も作業に入り、煩雑さを感じざるを得ません。そのハードルが下がってくればと願っています。

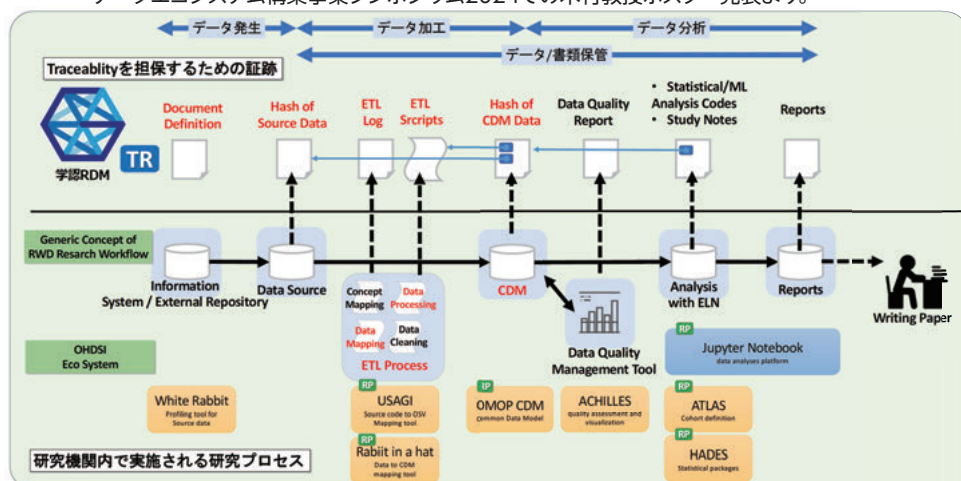
徳地 私のところでは、NII RDCに関しては途に就いたばかりです。というのも、今回のプロジェクトはいろいろな学問分野、いろいろな機関の先生に加わっていただいているのですが、それらの機関の多くが、学認に入っていない。どうやったら入れるの? というところからがスタートで、これがかなり大変でした。

結局、教授会に諮って、皆さんには京都大学の「連携研究員」になっていただいています。そこに漕ぎ着けるまでに結構時間を費やしてしまい、まだ「活用」と言える段階には至っていないのが現状です。今後進めていくにあたって、やはりもっと多くの方々にとって入りやすく、可能であれば研究者レベルでアプライできるものになってもらえればありがたいです。

小野 我々のところでは、まだ、大容量データの解析と、それをストレージ上で共有するというレベルにとどまっております。NII RDC上で全研究データを管理するというところまではたどり着いていないのが現状です。

我々が扱っているほど大量のデータ置き場として使用することが想定されているのかどうか。また大学内で使う分にはまったく問題ないと思うのですが、外部の方と連携をしたいという場合にどうすればよいのか。「自動自律実験」においては、メタデータの的なものも含め、人間を一切介

図2 「RWD※1研究の潮流となっているOHDSI※2のエコシステムを調査・類型化して、メタデータ定義や[GakuNin]RDMを利用した研究ベストプラクティクスの検討の基礎とする」研究データエコシステム構築事業シンポジウム2024での木村教授ポスター発表より。



在させずにデータの蓄積など全操作を行いたいのですが、こうした点もどのようにすればよいか。そのあたり、まだよく分からない点もあります。これも今後に向けての課題で、さらによく調べたり、NIIの方にお話を聞いたりしつつ進めていきたいと思っています。

これからの発展に向けて

——今後の抱負、展望について。

菊池 本事業で言うところの「研究データエコシステム」は、これからの研究活動の中で生み出される研究データが主な対象になっています。

ただ、人文系の立場からすると、これまで商業出版物として刊行された研究成果を、研究データとして流通することも視野に入れてほしいと思っています。もともと研究成果は紙の本にすることが規範でしたし、今なおそうなのですが、紙の本になっているがゆえに、そこに含まれる内容を研究データとして活用することが難しくなります。

いずれにせよ、人文系の研究は「研究データの管理」ということに関して、まだまだ馴染んでいないところがあります。NII RDCは、人文学でも利用

しやすい環境になるシステムを目指して頑張っていただけだと思います。

徳地 「研究データエコシステム構築」という、大上段からの展望というのは特になのですが、私たちが課題として提案した飛騨市の問題に関しては、まずは我々の持てるデータを有機的に関連させ、そう

した中から飛騨市の成り立ちのようなものが見えてくる、そんな活用の仕方を実現できていければと考えています。

木村 現時点では、まだNII RDCの中で、データ公開、データ管理、それから研究者の情報管理という要素が独立していて、統合化が図られていないと思います。そこを確実に進めていただきたいというのが第一の希望です。また医療分野では、個人情報の扱いの問題を含め、研究データの公開にはためらいがある場合が多いのです。これに関しては今まさに仮名加工に関するガイドラインが出てきたばかりですが、そうした動きに対し、研究者が個人的に対応するのは、や

はり負担が大きい。データを公開して大丈夫か、それを第三者機関が審査するといったプロセスもサービスの一環として設けていただけると、さらに利用も加速するのではと思います。

小野 冒頭に述べたように、現在、「自動自律実験」の分野では、ロボットなどを活用してデータを取得する部分が先行していて、海外の研究などでもそちらが注目を集めています。「研究データエコシステム構築事業」に参加させていただいて、我々がやらなければいけないことは、その先、自動自律のハードウェアの構築にとどまらず、データ側でどのように管理と活用を進め、それを通じて「本当に何か新しいもの、価値あるものを生み出す」というところまでつなげていくことだろうと思います。これは国内でも、世界全体でも現在足りていないですし、今後さらに重要性が増していくことです。そのためにはどのようなデータ基盤が必要なのか。それをNIIの皆さんとも一緒になって考え、作り上げていきたいと思っています。

※1：RWD (Real World Data) 診療行為に基づいて症状や調剤履歴等が記録され、匿名化された医療ビッグデータ。

※2：OHDSI (Observational Health Data Sciences and Informatics) 国際的ボランティアベースの医療ビッグデータ分析オープンサイエンスコミュニティ。



トランスディシプリナリ(TD)研究：市民交流会を開催

図3 「本研究で進めてきたこと ①地域デザインに資する材料収集②データプラットフォームの構築③『学認RDM』の活用と発展」研究データエコシステム構築事業シンポジウム2024での徳地教授ポスター発表より。

本研究課題では

①古典籍画像データからテキストデータを作成し、これらをTEI/XMLデータへと変換する。

②研究活動を通じて、古典籍TEI/XMLデータ作成ノウハウも蓄積、公開する。

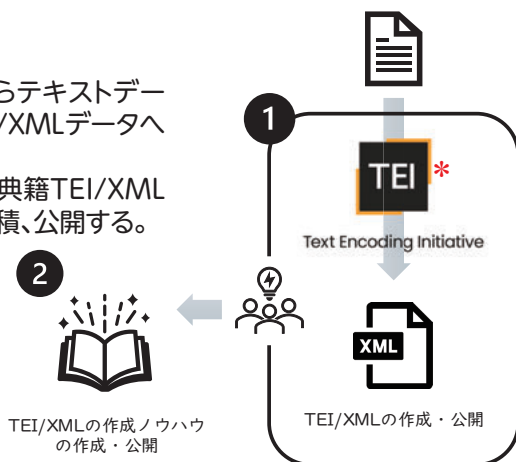


図4 国文学研究者に限らず、国内外の古典籍所蔵機関および異分野の研究者コミュニティに発信し、もってデータ駆動型人文学のための研究資源の確立を目指す「研究データエコシステム構築事業シンポジウム2024」での菊池准教授ポスター発表より。

* <https://tei-c.org/about/logos/>

ポリシーは起爆剤、 マネジメントは推進力 金沢大学の取り組み

2023年度、金沢大学はNIIの「研究データ管理スタートアップ支援事業」の地域拠点機関に採択され、他大学に支援を提供する立場となった。2024年度からは、福井大学、北陸先端科学技術大学院大学、金沢医科大学とともに「北陸研究データ基盤コンソーシアム」をリードする。日本の中でも先駆けとなる形で研究データのエコシステムを構築し、地域の特性を活かした研究開発を推進する姿勢と体制は、どのように生まれ、何を目指しているのだろうか。金沢大学の長井准教授に聞く。

データ管理のパイオニア精神

「金沢大学は、研究活動から生まれるデータを適切に管理することの必要性を、以前から強く意識していました。たとえば2020年、研究設備を大学間で戦略的に共用して研究活動を推進するコアファシリティプログラムに採択された時、データについても適切な管理を行う必要性を認識し、GakuNin RDMを導入し試行を始めました。最初からコアファシリティ連携を進めていたのは、金沢大学を含め2機関だけだったと思います」と語る、金沢大学先端科学・社会共創推進機構特任准教授の長井圭治氏。化学者として新素材・レーザー核融合など多様な分野での研究経験、分野や立場を超えたコミュニケーションの難しさ、乗り越えて協力する喜び、そして異文化融合のパワーで成果を生み出してきた経験がある。半導体リソグラフィの光源で数多くの業績を上げたのち、金沢大学に2021年着任。現在は研究の支援・推進を行うURAとして活躍する。文

科省「コアファシリティ構築支援プログラム」には、5年プロジェクトの2年目から加わった。

「データを適切に管理して利用するために、NIIが提供している研究データ管理基盤 GakuNin RDM を導入して試行を始め、2022年に『金沢大学学術データマネジメントポリシー』を策定していました。それが、2023年からのNII『研究データ管理スタートアップ支援事業(以下、スタートアップ事業)』につながったと考えています」(長井)

金沢大学とNIIの共同研究には十数年の歴史があり、数多くの成果を生み出してきた。その一つは、学術組織間でのデジタル分散共有システムである。全国の大学で機関リポジトリの整備が行われていた時期、金沢大学は次の一歩を踏み出していた。

北陸の、そして全国の「手本」を目指して

研究データエコシステムの構築と推進にあたって重要なステップを踏

長井 圭治

NAGAI, Keiji

金沢大学

先端科学・社会共創推進機構

特任准教授/URA

み出すために、背景があった。

「まず金沢大学の『未来知実証センター』で、データを公開しようという動きがありました。さらに、『コアファシリティプロジェクト』から、データサイエンスの推進やデータ利活用への方向性が出てきました。この2つの大きな背景があったところに、スタートアップ事業が始まりました。(図1)それが重要なポイントだったと思います。目標は北陸の手本、できれば全国の手本になることです」(長井)

研究を支援するICTシステムに関して、金沢大学には60年間の取り組みの蓄積がある。研究を推進するために必要不可欠な多様な技術基盤を提供する「総合技術部」があり、研究および教育に必要不可欠なICTシステムを担う「学術メディア創成センター」があり、NIIをはじめとする外部の研究機関との協力関係も蓄積されている。さらに、長井氏が所属する「先端科学・社会共創推進機構」があり、研究環境づくりや研究活動の支援に取り組んでいる。

「最初から研究推進と情報システムが連携していたことは、強みだったと思います。その中で、自然に生まれてきたとも言えます」(長井)

担当者や責任者同士の対話も活発で「お互いに、実効的な話をする事ができており、それも強み」ということだ。しかし大学の多様な部局には、それぞれ固有の文化や考え方があり、温度差やニーズの違いがある。それらは、障壁にならなかったのだろうか？

「それは、どの大学にもあることだと思います。私自身は多くの分野を経験しているわけではありませんが、極端に違う分野を経験してきました。それが役に立っているとは思いますが」(長井)

たとえば事故リスクと隣り合わせの有機化学と、発想や多様な議論が重要な理論研究では、「研究室のコアタイムは、どの程度守られるべきか」という点でも温度差がある。さらに、医学部と大学病院には、人命にかかわる日常がある。

「私自身は医学部との共同研究の経験はありませんが、金沢大学の現在の学長は医学部出身です。医学部は設備の共同利用が進んでおり、ルールも精緻です。理学部は、そこまで細かいルールを設けず、自由な発想を大切にします」(長井)

一方が正解で、もう一方が誤りというわけではない。医学部のルールを適用された部外者は、「厳しすぎる」「自由に情報を利用したい」といった感覚を抱きがちだ。しかし、その厳重さには理由がある。

「医学部のルールを他に展開するのは、それほど難しくありません。でも、逆は困難です。ルールやポリシーを策定するときは、『これを医学部で

動かしてみたら、どうなるか』という意識を持つと、成功しやすいのではないかと思います」(長井)

大学の責任を明確にした

データマネジメントポリシーへの道

金沢大学は、2022年3月に「金沢大学学術データマネジメントポリシー」を策定し、2024年7月には実効的なデータ公開への対応を中心とした改訂を行った。その中心は「データの取り扱いの責任は大学にある」という明確な規定、論文と学会発表以外でデータを公開するにあたっての判断を行う機関の一元化、違法なデータ取得を排除するためのデータソース把握。かなり踏み込んだ内容となっているが、あくまでも「ポリシーは、憲法です」と長井氏は語る。肉付けや具体化は、各部局や各部署の役割である。そもそも、ポリシーは、何のために必要なのか？

「オープンサイエンス推進のためには、研究データの適切なマネジメントが不可欠、そのためにはマネジメントのポリシーが必要です。『第6期科学技術・イノベーション基本計画』でも、各大学の研究データポリシー策定やデータマネジメントを目標として掲げています」(長井)

しかし「策定しなくてはならないから、策定した」というポリシーでは意味がない。

「ポリシーは、研究データの管理や利活用を推進するルールとして策定すべきものです。それは、実のあるオープンサイエンスを実現するための起爆剤にもなり得ます」(長井)

長井氏が金沢大学に着任した時期は、ポリシー策定のためのワーキンググループが動き始めていた。まずは、先行していた京大や名大と概ね同様の内容に教育データを含める形で、「学術データマネジメントポリシー」として策定された。

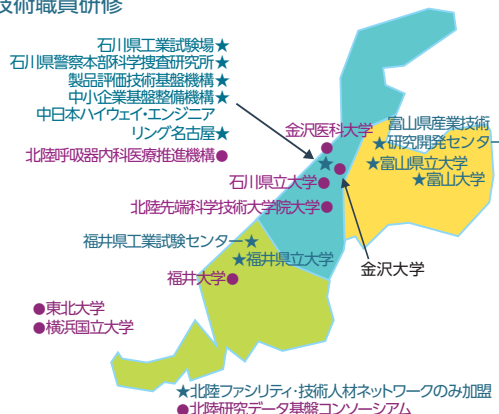
「先行例を参考に、なるべくシンプルにポリシーを作ることに集中しました。最初から『あれもこれも』と欲張ってしまうと、合意形成が難しくなります」(長井)

運用を開始すると、「教育データ」を含めたからこそその課題が浮かび上がった。データには、いったん登録した後での変更があり得る。必要な対応は、一様ではない。そこで、2023年度から改訂作業が始まった。

地域連携の戦略

北陸ファシリティ・技術人材ネットワーク(コアファシリティ事業)

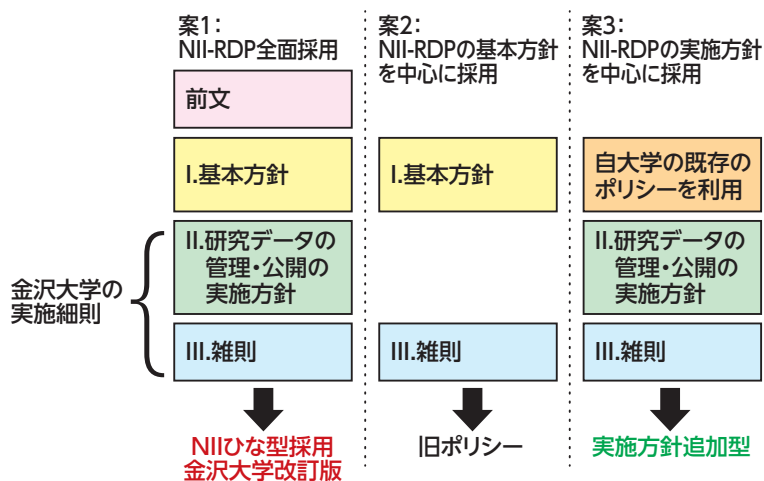
ファシリティデータベース、技術者データベース、技術職員研修



北陸研究データ基盤コンソーシアム(スタートアップ事業)

研究データポリシー策定支援、オープンアクセス加速化事業の成果共有、研究データエコシステム構築に関する活動FD研修

図1 北陸研究データ基盤コンソーシアムは、金沢大学を中心とし、学術データの管理・公開・利活用の支援を目的として、各機関のルールづくりやそれを支えるシステムの開発や展開を進める活動を進めている。
<https://dri.w3.kanazawa-u.ac.jp/consortium/about.html> および長井氏資料を元に作成



- ポリシーは憲法
- 改訂に関してはNIIひな型を取り入れつつ法的観点から整備
 - ①大学の責任（主語を「大学（本学は）」で統一）
 - ②データ公開の観点
 - 学術論文、学会発表以外は「分譲委員会」に諮問
 - ③データソースの把握（違法なデータ取得を排除）
 - 転出、退職、共同研究などのデータのやり取りを規定
- 今後詳細は「実施細則」で規定
「ガイドライン」はポリシーと実施細則の説明

図2 金沢大学の研究データポリシー策定
(長井准教授資料を元に作成)

※NII-RDP=研究データ管理・公開ポリシー
<https://rcos.nii.ac.jp/service/datapolicy/>

「ポリシーは、組織の中に向かって重要なルールですが、同時に大学の対外的な宣言でもあります。外部の方が読んで一義的に取れるように、主語は全て『本学は』とするなど、弁護士さんに指導を受け、現在の形になっています。(図2) NIIのポリシーひな型がそのまま使えるかと言えば、そう単純ではありませんが、チェックリストのように非常に役に立ちました」(長井)

むろん、学内での試用や結果の検討も容易ではない。

「公開するデータを持っている研究者にケーススタディしてもらい、事務職員を含めて多様な立場の人々で議論しているところです。データマネジメント自体が新しい概念なので理解されにくかったり、全員が同じ意味で理解できる用語を探す努力を互いに重ねたり、そのポリシーを実施した結果に関する共通のイメージを持てるようにすることに苦心したりもしました。合意に至りにくいことは、あえてポリシーに含めない判断もしました。ポリシーには、絶対に変えないことを書くべきです」(長井)

細かいこと、状況次第で変わることは、実施細則で規定すればよい。しばしば変更しなくてはならないようでは、大学への信頼にもかかわる。ま

さに、ポリシーは憲法だ。

金沢だから、北陸だからその可能性と課題

2024年1月には能登半島地震、9月には『令和6年奥能登豪雨』に見舞われた。被災した方々の痛みを思いつつ、金沢に残る、古い資料や文化財をデータとしても後世に伝えるため、金沢大学も奔走。その中で「情報が散逸しないうちにGakuNin RDMに保存せよと理事から求められました。しかし、そこで浮かび上がった課題の1つ目は、使用にはノウハウがある程度必要であること。2つ目は、加盟機関が限られ、私立大学はほとんど未加入、高校の先生などはアクセスできないこと。3つ目は、上記から緊急性のある事案に使いづらいこと、です」(長井)

これらは、NIIへの宿題でもある。「ルールも仕組みも始まったばかり。まだまだ定常運転という段階ではなく、重要な事例が生まれるところで、一つ一つの確に、できればスピーディに対応したいと思っているところです」(長井)

これから研究データポリシー策定に取り組む組織へのヒント

「最初は、なんでもかんでも取り入れるというのではなく、まずは全学的な合意形成を試みるということ。小

さなことでいいので、データを公開しようというポリシーを作る、そこからですね。同時にいろいろテクニカルなことは少しずつ進めなくてはなりませんが、『学認って何?』というところからのスタート支援は、NIIのサポートなしにはできません。ただ、体験しないと分からないところも多いので、本当に何も知らない人に向けたメニューも必要かもしれません。入り口は重要です。

その次にポリシーを改訂しようという段階になったら、今後いろいろな大学から先行例が出されると思いますので、それぞれ比べて、使いやすいところを真似するというのが現実的だと思います」(長井)

ポリシー策定に関する懸念もある。「一つ危惧しているのは、大学が責任を持たずに、部局に任せるとするのがデフォルトのようにになっている点です。組織改編などの場合に、責任の所在が不明確になるので、大学がある程度、責任を持ったポリシーを作る必要があるのではないかと思います」(長井)

大学は、研究データポリシー策定を考えた時、長井氏のような先人が活動する、地域コンソーシアムなどを覗いてみることで、さらなるヒントが得られるかもしれない。

NII ニュース・トピックス一覧

NEWS TOPICS

期間: 2024/8/1 (木) から
2024/12/31 (水) まで



各ニュースの詳細は
オンラインでご覧になれます。

www.nii.ac.jp/news/2024

Facebook

[www.facebook.com/
jouhouken/](https://www.facebook.com/jouhouken/)

X (旧Twitter)

x.com/jouhouken

YouTube (音が出ます)

[www.youtube.com/user/
jyouhougaku](https://www.youtube.com/user/jyouhougaku)

情報犬ビットくん

X (旧Twitter)

x.com/NII_Bit

NII Today

ご意見募集中

www.nii.ac.jp/today/iken

メルマガ

購読随時受付中

www.nii.ac.jp/mail/form

NII Todayはオンラインでバックナンバーをご覧になれます。

www.nii.ac.jp/today

冊子版は、学校図書室、研究機関など団体への配本を受け付けています。お申し込み、配本停止は、kouhou@nii.ac.jp 広報へメールでご連絡ください。個人への送付は行っておりませんが、最新号の冊子は東京・千代田区一ツ橋の学術総合センタービル1階に配架されており、ご自由にお持ちになれます。

ニュースリリース

NEWS RELEASE

2024

- 12/24 完全にオープンな約1,720億パラメータ(GPT-3級)の大規模言語モデル「llm-jp-3-172b-instruct3」を一般公開 —GPT-3.5を超える性能を達成
- 11/28 CTCとNII、学術機関におけるデジタル資格証明の標準化に向けた共同研究を推進
—学位・学修歴・学生証のデジタル化に向けたガバナンスと技術標準の策定へ—
- 11/1 国立情報学研究所の公式キャラクター「情報犬 ビットくん」LINE スタンプ第6弾を発売
- 10/21 CiNii Research 自動翻訳機能 試行版の提供開始
—日本語が理解できない利用者でも日本語論文の検索が可能に—
- 10/16 富士通と産学組織が9者で共創し、世界初の「偽情報対策プラットフォーム」の構築を開始
- 9/17 約1,720億パラメータ(GPT-3級)の大規模言語モデルのフルスクラッチ学習を行い、プレビュー版「LLM-jp-3 172B beta1」を公開
—学習データを含めすべてオープンにしたモデルとしては世界最大—
- 9/2 情報学の技術を使って化学の宝探しをしてみませんか?
メタンからエタンを効率的に生成する触媒を機械学習の手法で探すコンペティションを開催

受賞

AWARDS

2024

- 12/5 新井 紀子教授(情報社会相関研究系)らの論文が一般社団法人言語処理学会の言語処理学会30周年記念論文賞を受賞
- 12/4 蓮尾 一郎教授(アーキテクチャ科学研究系)らによる論文が ICTAC 2024 で Best Paper Award を受賞
- 11/3 根岸 正光 名誉教授が瑞宝中綬章を受章
- 10/2 飯野 なみ 元NII特任研究員(情報学プリンシプル研究系)らによる論文が情報処理学会論文誌ジャーナルの特選論文に選出
- 9/26 井上 克巳教授(情報学プリンシプル研究系)らによるNII 共同研究論文が LPNMR 2024 で最優秀学生論文賞を受賞
- 9/11 「人文科学研究助成」採択の坊農 真弓准教授(情報社会相関研究系)に三菱財団が贈呈書を授与
- 9/5 小林 泰介 助教(情報学プリンシプル研究系)による論文が日本ロボット学会 第38回 学会誌論文賞を受賞
- 8/15 五島 正裕 教授(アーキテクチャ科学研究系)らが xSIG 2024において Outstanding Research Awardを受賞
- 8/15 内山 一秀さん(総合研究大学院大学、五島研究室)が xSIG 2024において Poster Awardを受賞
- 8/6 坂井田 瑠衣さん(元NII特任研究員、公立はこだて未来大学)、坊農 真弓 准教授(情報社会相関研究系)の論文が日本質的心理学会 優秀現場研究報告賞を受賞
- 8/2 小川 潤 特任研究員(コンテンツ科学研究系・北本研究室)の論文が情報処理学会 2024年度 山下記念研究賞を受賞



←巻頭写真:福島県 JR浪江駅目の福島国際研究教育機構(F-REI)建設予定地にて撮影。ここは福島第一原発から5キロほどの地点。一部地域では避難指示が継続しているものの、2024年現在、浪江町全体では約2千名が居住している。江村克己氏が理事を務めるF-REIは、福島復興再生特別措置法に基づき2023年に設立、科学技術力・産業競争力をリードする「創造的復興の中核拠点」を目指し、この地で活動を始めた。

Essay

データ利活用研究コミュニティの創成に向けて

田浦 健次郎

TAURA, Kenjiro

東京大学執行役・副学長
同大学院情報理工学系研究科 教授

研究データエコシステム構築事業の中で「ユースケース創出」という事業を実施しています。そこでは研究データの整備や利活用に主眼を置いたプロジェクトを公募・審査をし、採択された課題を実施するのに必要なサポートを行っています。2025年1月現在までに33件の課題を採択しています。公募情報は <https://www.nii.ac.jp/creded/use-cases.html> に、これまでおよび現在、実施している課題の一覧は https://www.nii.ac.jp/creded/creded_result.html にありますのでぜひご一読ください。

採択された課題には、GakuNin RDMやmdxなどの、データを蓄積、共有、活用する基盤を積極的に活用していただき、課題を推進していただけます。また、口頭発表やポスター発表、パネルディスカッションを行うシンポジウムを開催し、ユースケース課題実施者から、事業全体へのフィードバックやユースケース課題実施者同士の意見交換、情報交換を行っています。ユースケース課題のSlackワークスペースも設営しており、GakuNin RDMやmdxなどの基盤に関する質問を随時受け付けています。

GakuNin RDMやmdxなどの、データを蓄積～共有～活用する基盤を活

用する研究を支援・推進・創出

情報技術や情報基盤を活用した研究は従来から多くの分野で重要でしたが、AIの著しい進展も手伝って、データを集積、整備しそれを検索可能にすること、そのデータを活用することの重要性がさらに多くの分野で増えています。本事業では、GakuNin RDMやmdxなどの全国の研究者に提供されている研究基盤を積極的に活用することにより、研究データの蓄積、整備から、共同研究者との共有、AIなど大きな計算資源を要するものを含めた大規模データ処理を一貫した環境で推進することを支援しています。

全国的なデータ基盤の永続と高度化、それによる研究環境、研究力の向上

そのような基盤を利用した研究の進め方は、本ユースケース創出事業の課題であるか否かを問わず、研究の進め方の標準として促進されるべきものでしょう。そのような基盤が常に最新の設備に更新され続け、全国的に(所属機関を問わず)、迅速に、廉価に利用可能であることが当たり前になれば、各研究者や機関、または分野ごとに個別に研究環境を構築するコストも時間も削減でき、研究力の向上につながります。また、データ基盤、活用基盤と言っても情報基盤に課す要求はさまざまで、実験機器との連携、

セキュリティの向上、大容量データの長期保存、耐故障性や耐災害性などがあります。本ユースケースのような事業を通じた成果により、そのような基盤の必要性が広く理解され、分野ごとに異なるさまざまな要求を取り入れた次世代基盤を作っていく流れにつながることを目指しています。

情報学・さまざまな分野をつなぐ学際的な研究

もともと情報学という分野は、多くの分野で利用可能なツール(文字通りのツールという意味にとどまらず分野を超えて使われる共通の計算手法やアルゴリズム)を提供している分野であり、他の分野との結び付きが広く存在します。本事業では、その結び付きを広げ、情報学と他の分野、または情報学を活用する複数の分野をつないだ学際的な研究を生み出すことを目指しています。

終わりに

このユースケース創出事業では、常時応募を可能とし、隔月で審査を行っています。また、一度の応募で2年度にわたる申請が可能であるなど、研究を効果的に支援するための工夫を行っています。趣旨に合った研究テーマをお持ちの方はぜひ応募をご検討ください。もちろんGakuNin RDMやmdxの利用は本事業に応募しなくても可能ですのでそちらもご検討ください。

情報から知を紡ぎだす。



国立情報学研究所ニュース:NII Today 第104号 令和7(2025)年1月
発行: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
〒101-8430 東京都千代田区一ツ橋2丁目1番2号 学術総合センター

本誌についてのお問い合わせ:総務部企画課 広報チーム
EMAIL: kouhou@nii.ac.jp

発行人:黒橋 禎夫

編集委員長:越前 功 監修:安浦 寛人 アドバイザー:中野 恵一

編集委員:池畑 諭、金子 めぐみ、込山 悠介、水野 貴之、竹房 あつ子 (ABC順)

編集/AD:岸本 治恵 制作・印刷:日精ビーアール

表紙イラスト:市村 譲

この製品は、適切に管理されたFSC®認証林、再生資源およびその他の管理された供給源からの原材料で作られています。VOC(揮発性有機化合物)石油系溶剤を含まないインキを使用し、有害な廃液を出さない「水なし印刷」で印刷しています。印刷に使用する電力の13.8%はバイオマス発電によるグリーン電力を利用しています。「紙」「インキ」「加工資材(綴じ具)」はいずれも、紙・板紙へのリサイクルにおいて阻害にならないリサイクルランクAの材料を使用しています。



この印刷物は、印刷用の紙へリサイクルできます。