

AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ（第5回） 配付資料

https://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu29/006/siryo/mext_00007.html

記事ページ本文

現在位置

トップ

>

政策・審議会

>

審議会情報

>

科学技術・学術審議会

>

情報委員会

>

AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ

> AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ（第5回）配付資料

AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ（第5回）配付資料

1. 日時

令和8年5月22日（金曜日）16時00分～18時00分

2. 場所

文部科学省東館17階局4会議室 及び オンラインのハイブリッド形式

3. 議題

AI for Scienceを支える研究データの管理・利活用及び流通の在り方について

AI for Scienceを支える研究データの管理・利活用及び流通の在り方ワーキンググループとりまとめ素案について

その他

4. 配付資料

【資料1】 AI for Science を支える研究データの管理・利活用と流通の在り方ワーキンググループとりまとめに向けた整理（第4回審議追加版）(PDF:4.5MB)

【資料2】 ライフサイエンス分野のデータ基盤 (PDF:3.5MB)

【資料3】 マテリアル分野におけるデータ基盤蓄積から活用に向けた課題 (PDF:2.0MB)

【資料4】 大型放射光施設SPring-8の研究データ基盤とAI for Scienceに向けた取り組み (PDF:2.2MB)

【資料5】 AI for Science実現のための基盤整備について (PDF:1.6MB)

【資料6】 AI for Science を支える研究データの管理・利活用及び流通の在り方ワーキンググループ今後の進め方及びとりまとめの方向性について（案）(PDF:490KB)

【参考資料1】 科学技術・学術審議会 情報委員会 AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ運営規則 (PDF:208KB)

【参考資料2】 海外公的機関におけるAI for Scienceを見据えた研究インフラ（プラットフォーム）の動向 (PDF:2.2MB)

お問合せ先

研究振興局参事官（情報担当）付学術基盤整備室

PDF形式のファイルを御覧いただく場合には、Adobe Acrobat Readerが必要な場合があります。

Adobe Acrobat Readerは開発元のWebページにて、無償でダウンロード可能です。

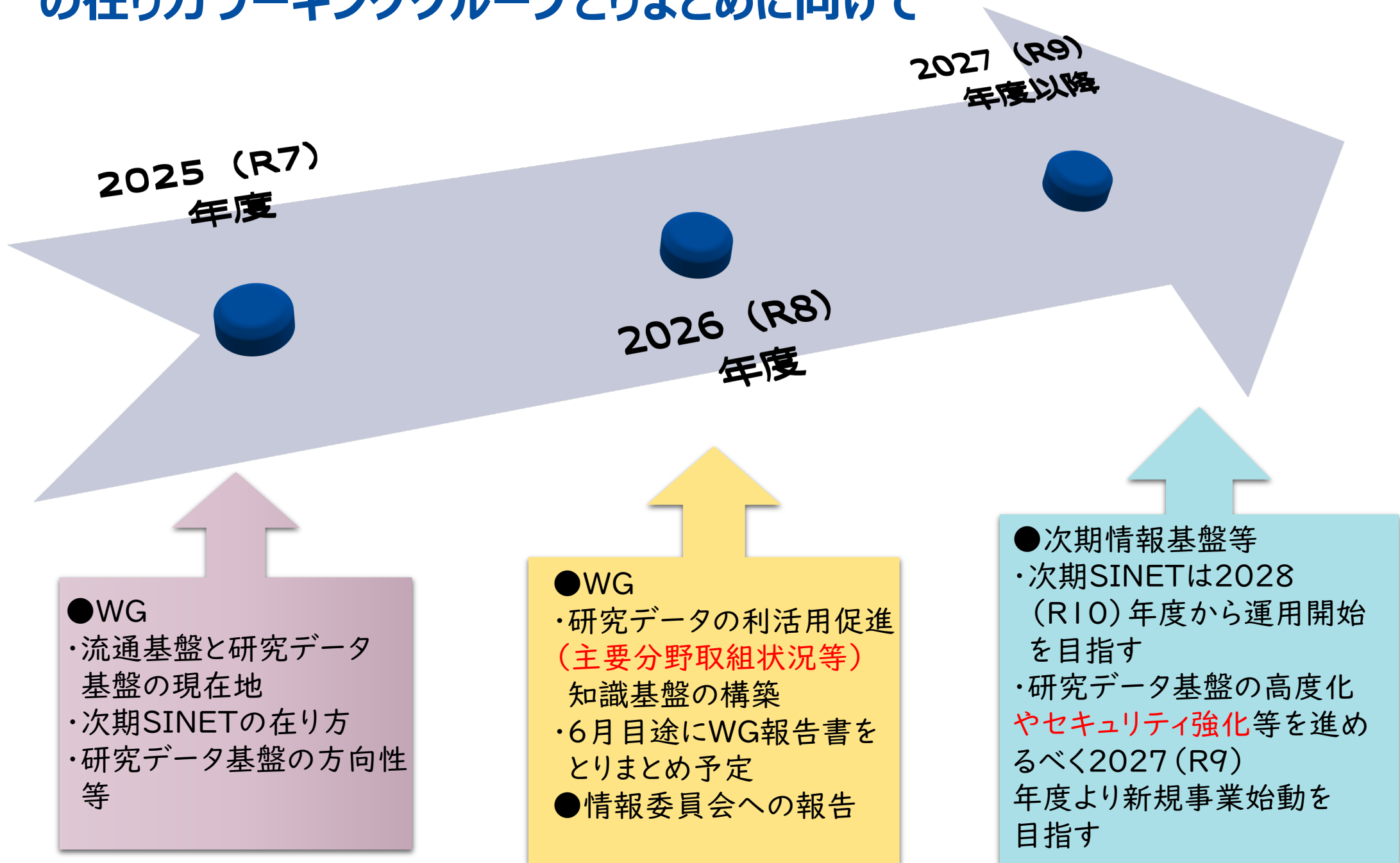
（研究振興局参事官（情報担当）付学術基盤整備室）

[ページの先頭に戻る](#)
[文部科学省ホームページトップへ](#)

AI for Science を支える 研究データの管理・利活用と流通の在り方 ワーキンググループとりまとめに向けた整理 (第4回審議追加版)

2026年5月22日

AI for Science を支える研究データの管理・利活用と流通 の在り方ワーキンググループとりまとめに向けて



※文部科学省におけるAI for Science 推進の動きとも適時連動

整理すべき内容（とりまとめ骨子）

○研究データの管理・利活用と流通を支える次世代情報基盤の必要性

- 背景（AI for Science 推進とオープンサイエンス進展）
- 政策の方向性（AI for Science の推進に向けた基本的な戦略方針、第 7 期科学技術・イノベーション基本計画等）

○学術研究プラットフォームの構築・実施状況

- 情報流通基盤（SINET 6、NII-SOCS）の現状
- 研究データ基盤（NII RDC）の現状

○国内外の動向

○ユーザーニーズ、社会からの要請

○AI for Science の推進に向けた基本的な戦略方針への対応

○次世代情報基盤整備の方向性

- 付加価値の最大化を実現するネットワークの整備
- 研究データ管理・利活用の活性化、基盤高度化
- 計算資源との有機的な接続・連携、認証強化

【流通基盤 SINET】

- 現行SINET（SINET6）の活用状況
- 次期SINETに向けたニーズの把握
- 最新の技術動向
- 海外の情報通信ネットワークの動向
- サイバーセキュリティへの対応

【研究データ基盤 NII RDC】

- NII RDCの活用状況（高度化の効果、ムーンショット等での先行利用等）
- 利用者（研究者）ニーズの把握
- 最新の技術動向（認証、AIの利活用含む）
- 海外の研究データ基盤の動向（欧州：EOSC等）
- 計算資源との有機的な接続（解析・保存等、mdx等との接続含む）
- AI for Scienceやオープンサイエンスの推進・加速対応のための機能高度化の要件
- データ管理・利活用に資する研究支援等人材の育成

etc..

これまでのワーキンググループ(WG)での主なポイント等

- ・第1回開催 2025.12.24
- ・第2回開催 2026.2.13
- ・第3回開催 2026.3.26
- ・第4回開催 2026.4.24

第1回WGでの主なポイント〈委員コメント等〉

・NIIがこれまで構築を進めてきた「学術研究プラットフォーム」の現在地と今後の展開についてNIIより説明

【流通基盤 SINET】

- 民間企業の利用拡大とネットワークの役割として、研究データの利活用先として民間企業を含める方向性は理解できるが、「誰でも使えるようにする」モデルは現実的ではない。ネットワーク・基盤の利用主体をどう線引きするのかが重要
- AI for Science時代のデータ流通要件として、AIによって研究プロセス自体が破壊的に変わる可能性があり、データ流通基盤も「従来型の増強」では不十分となる可能性を視野に入れるべき。
- 実験側（ロボット化・自律化）と解析側を結ぶネットワークが重要。

【研究データ基盤 NII RDC】

- 共通基盤化は、研究プロセスの標準化効果が期待される。JAIRO Cloudの例は、各大学が個別に運営するよりも、共通基盤化により「自動的に標準化が進む」好事例である。
- 研究データ基盤は単なるITではなく、研究者の作業負担軽減、時間創出（＝研究の質向上）に直結する。
- コストと集約の視点として、個別大学・研究者がそれぞれ調達・交渉する構造は非効率であり、調達・管理・交渉の集約化余地が大きい。

セキュリティ

- 個人情報・医療データ・社会データ等の機微データを扱うには「安心して置ける／流せる」基盤が不可欠。倫理・法対応だけでなく、技術基盤としての信頼性をNIIが担う意義は大きい。
- データの性質（質・機微性）と利用者属性に応じた設計が必要

Ai for Scienceへの対応

- AI for Scienceを推進するうえで、AIを「解析支援」に矮小化すべきでなく、実験自律化、仮説生成、実験プロセスへのフィードバックまで含む概念としてとらえるべき。そのうえで、学術研究プラットフォームは、どこまでカバーするのか、マテリアルなど、分野固有のデータ基盤を持つ外部のプラットフォームとどうつながるのかなど、他分野・他機関の基盤とどう連携するのかを明確化すべき。

研究データの扱い等

- 個人情報保護法の今後の改正により、AI高度化社会課題解決に資する場合、同意によらないデータ利用が可能になる領域が広がる。学術研究用途は、一般的な民間利用よりも社会的受容性が高いと考えられる。そのため、「個人データは使わない」という前提ではなく、どうすれば適切に使えるかという方向で議論すべき

第2回WGでの主なポイント（流通基盤、セキュリティ） <NIIの構想ポイント>

- ・SINETの高度化、セキュリティ強化に向けた構想についてNIIより説明
- ・AI for Science推進に向けた基本的な戦略方針の検討状況について文部科学省より説明

【流通基盤 SINET】

機能強化（速度）

- 今後もトラフィック流量の増加が予想されるため、次期SINETでは、新しい800ギガ技術を用いて速度を上げていく必要がある。
- さらなる需要増も予測されるため、適切なタイミングで800ギガから1.6テラへ増速する2段階構成での検討も進めている。

機能強化（安定性・信頼性）

- 次期SINETでは、SINET6と同様に、3層アーキテクチャを用いることでデータセンター間の最短距離通信を実現する。
- IOWN技術等を取り込み、新たに光パスレイヤーの切替えを実現する。光伝送レイヤーのファイバーが切れたときに、別のファイバーを用いて通信を再開することができるため、信頼性向上に繋がる。
- 現在1台構成のルーターを2台とすることで信頼性向上を図る。
- VPNサービスを拡充し、一つの経路しか使えないアクト・スタンバイ型から2つの経路を使用可能にする。

モバイルの強化

- 携帯電波が届かないエリアでの研究活動を支援するため、衛星通信や飛行船（非地上波通信）を介した通信と組み合わせ、より広い屋外での教育研究を支援する。

海外（米国・欧州）

- 次期運用期間途中に、現在400ギガで接続している米国・欧州の回線を倍の800ギガに増強することを検討。
- 新たな海底ケーブル敷設においては、地政学的な観点からより安定的なケーブルを選定し構築することを予定。

スケジュール

- 今後2年間で調達・設計・構築を進め、2027年度の最終クォーターで移行作業を行い、2028年4月から開始する。

セキュリティ

- インターネットとVPNの両方におけるセキュリティを確保するため、新たなネットワーク制御技術を取入れ、論理分割を行うことも検討。
- NII-SOCSによる攻撃の検知、DDoS Mitigationによる攻撃を受けた際のトラフィック防御機能を組み合わせ、一体的なサービスを提供する。
- セキュリティ用の装置を複数大学で共用する仕組みを用意することで、設備費用の低下を支援する。
- NII-SOCSのサイバー攻撃検知・解析の対象を公私立大学に拡大することを検討。
- 複数の大学でクラウド型のファイアウォールを共同利用する仕組みを検討。
- 大学の認証システムに障害が発生した際に、認証基盤のホスティングを代行する機能について検討。

第2回WGでの主なポイント（流通基盤、セキュリティ）〈委員コメント等〉

【流通基盤 SINET】

高度化

- 学術網としてユーザーが安全安心に使えるようなサービスを提供することに加え、世界最先端の技術やサービスを提供できる場であってほしい。
- ネットワーク構成検討にあたり、必要な容量という観点だけでなく、世界最先端の研究実現を念頭に置いて検討してほしい。
- AI for Scienceで実験の自動化等が進行すると、データ容量の伸びが桁違いに増えると予想される。莫大なデータを持つコアファシリティ間を繋ぐように広がっていくと、データ容量の伸びは加速する。今後のSINETについて検討する上で、どこまでデータ容量が伸びるのか予想をするのが難しい。
- 今後学生がAIを本格的に利用すると、現在の見通しよりもトラフィックの伸びが高くなる可能性がある。初等中等教育機関において、今まで以上にSINETを利用するようになると、より一層伸びが加速することが見込まれるのではないか。

ネットワーク構築

- Spring-8やNanoTerasuのような放射光施設をはじめとした大型プロジェクトでは、非常に多くのデータが蓄積されるので、その点に留意しネットワーク構成を検討すべき。
- ネットワークの整備計画は、大型実験施設の整備計画やコンピューティング基盤の整備と連携して考えた方が良い。

ネットワークの運用・開発を行う人材不足

- ネットワークの運用・開発に関わる人材が減少しているので、そうした人材を育成し、現場のマネジメントができる環境を維持してほしい。

セキュリティ

- 安全保障の観点からネットワークそのものの機能や運用にも注意を払わなければならない。セキュリティを適切に担保できる装置を選んでいくことも重要。
- 今後はファイアウォールの中で守るというよりは、ゼロトラストに対応することが必要となってくる。大学のゼロトラストに対する取組に対し、支援等ができないか。
- AIの利活用によって生じる、サイバーセキュリティや情報インシデント等への対策についても検討すべき。

認知度向上

- SINETやNII RDC等は日本の基盤的なインフラだと思うので、より一層一般の研究者にリーチするための活動を行うべき。
- NIIと各大学担当者との連携を強化し、SINETの認知度向上に努めた方が良いのではないか。

第3回WGでの主なポイント（研究データ基盤、認証） <NIIの構想ポイント>

- ・研究データ基盤の高度化、認証強化に向けた構想についてNIIより説明
- ・AI for Science推進に向けた基本的な戦略方針の検討状況について文部科学省より説明

【研究データ基盤 NII RDC】

高度化

- NII RDCを、研究データ管理における我が国の中核基盤として継続的に高度化することが重要であり、単なる保存基盤ではなく、研究セキュリティ・研究公正・オープンサイエンスの双方を支える基盤として機能させる必要がある。
- DMPに基づいて研究データ管理を研究中から支援できる仕組みを整備することが今後の方向性。
- NII RDCはこれまで「データができた後」からの関与が中心であり、実験装置やコアファシリティとの接続が弱かったという認識。実験装置やコアファシリティとNII RDCを接続し、データ発生源から基盤までをシームレスにつなぐ取組が重要であり、EPOCHや自動化ラボ等の新規事業と連携し、データ創出段階からRDCを活用できるようにする必要がある。
- メタデータ作成は研究者にとって大きな負担であり、AIによるメタデータ補助入力が課題。実験メモや電子ラボノート等を活用し、AIでメタデータを生成・整理する取組を推進する必要がある。今後、メタデータだけでなく、収集・解析・コーディング段階でもAIを活用し、研究者の負担を軽減する機能を提供する。
- 実験装置・ストレージ・計算資源をつなぎ、データ収集、解析、実験管理を支援するデータ駆動基盤の構築が必要。

次世代認証基盤の構築・フェデレーション強化

- 認証基盤の抜本的拡張・次世代化が重要であり、大学IdPに依存した現行のGakuNinモデルでは、共同研究者や未参加機関の研究者を十分に包含できない。研究グループ（バーチャルオーガニゼーション）単位での認証・認可や貴重なデータ・計算資源に対応するための高認証強度（多要素認証等）の普及が必要。
- NII RDC単体ではなく、mdxや国内外の計算資源と連携することが不可欠であり、特定機関だけでなく、幅広い研究者が計算資源を利用できるフェデレーション型の仕組みを検討する必要がある。さらに、欧州（EOSC、EuroHPC）を参考にしつつ、日本型の研究リソース連携モデルを構築する必要性を認識。

【その他・総論】

地域コンソーシアムの拡大

- 各大学・研究機関が個別対応するのではなく、地域コンソーシアムにより課題を共通化し、効率的に議論・対応することが重要であり、地域コンソーシアムを通じて、研究データポリシー、オープンサイエンスへの対応、データ主権、研究セキュリティに組織として対応していく必要がある。
- NIIは基盤提供者として、地域コンソーシアムを支援する役割を担う。

人材育成

- 研究データ管理・オープンサイエンスに対応できる人材育成が不可欠。教材・研修コンテンツを拡充し、ファカルティ・若手研究者・URA等への展開を進める必要があるが、NII単独では限界があるため、地域コンソーシアム等と連携しながら普及を図ることが重要。

第3回WGでの主なポイント（研究データ基盤、認証） <委員コメント等>

【研究データ基盤 NII RDC】

AI for Science推進と研究データ基盤

- 研究データ基盤も、単にAIのためのデータ置き場になるのではなく、実験・現場・研究者にどう寄与するのが重要。多くの研究者に「自分は関係ない」と思われてしまった時点で、この取組は負けだと思ふ。AI for Scienceの動きを、リアルな研究をしている人たちが、自分たちの研究にどう取り入れるかが見えるようにしていく必要がある。
- ネットワークでデータを集めるところまでは、今回の説明でカバーされているように感じた。しかし実際には、いろいろな分野からデータを吸い上げること自体が容易ではない。研究データ基盤を考えるにあたっては、データを運ぶことデータを管理することだけでなく、そもそもデータをどう創出・収集するかという点にも、もっと重点を置く必要があるのではないか。
- AI研究においても、インターネット上のデータは枯渇しつつある今後は学術が持つ一次データの重要性が増す。医学系や病院のデータも、外には出せないが非常に価値が高い。セキュリティを確保した上で、そうしたデータも含めて研究に活かす設計が重要ではないか。

認証とアクセスの確保

- 人文系の資料が大量に眠っているが、それらの多くはまだデジタル化されていない。場合によっては、紙であるがゆえに朽ちてしまう可能性もある。人文社会系のデータは、データがそこに存在していても研究に使える形になっていない場合が多い。インターネット上のデータは枯渇しつつあると言われているが、今後は、学術研究が持っている一次データが、非常に重要になってくる。
- 医学系や病院のデータについては、外に出せないデータが多いが、非常に価値が高い。セキュリティに十分配慮した上で、そうしたデータも研究に活かせる環境を考えることが重要ではないか。

【総論】

NIIの役割

- NII自身は、計算機や計算サービスそのものは提供しておらず、スパコン、学内HPC、商用クラウド上の計算資源は、NIIの直接の提供対象ではない。これらは基本的に、大学の基盤センターや商用クラウド事業者が提供主体であり、NIIの役割は、計算資源を「うまくつないでいくこと」にある。単一の計算基盤を用意するのではなく、既存の計算資源を活かしそれらを研究者が使えるように結ぶことがNIIの立場。

必要な支援

- 単にツールやインフラを提供するだけでは不十分。新しい人が入ってきて、システムを使おうとしたとき、問題が起きたとき、分からないことが出てきたときなど、相談できる人的サポートがあると活用が進むのではないか。また、専門家でも状況が分からなくなることがあるため、エキスパートやエンジニアが支援する仕組みがあるとよいのではないか。

第4回WGでの主なポイント（知識基盤の構築） <NIIの構想>

・知識基盤の構築に向けた構想についてNIIより説明

【知識基盤の構築】

現状課題・方向性

- 知識基盤は、学術研究プラットフォームの最上位に位置し、AI（LLM等）により研究データを知識へ変換する中核機能として、研究データ基盤の上にAI機能を統合し、データから知識を創出する基盤。
- NIIは、国産のオープンLLM（LLM-jp）を開発し、現在はQwen等と同等性能のモデル（LLM-jp-4）を実現しているが、学術研究においては再現性・透明性の確保が不可欠であり、オープンなAI基盤により、データ主権及び研究の自由度を確保することが必要。
- 国際的な知見共有により効率的な開発を図りつつ、国際連携の下でAI基盤の高度化を推進し、自律性と国際協調を両立したモデル構築が必要。
- 現在、研究者は特定の外部AIサービスに依存しており、コスト負担やデータ制約・再現性の課題が顕在化。さらに研究過程における未公開データや医療・社会データ等の機微データの利活用には制限があり、十分な活用ができていない。
- 論文査読や研究計画の検討など、機密性の高い研究活動においてAIの活用ニーズが存在するにもかかわらず、適切に利用できる環境が未整備。
- 分野特化型のAIモデル構築や追加学習については一定のニーズが存在するものの、必要な環境や知見が不足しており、広く普及できていない。
- AIの高度な活用には、プロンプト設計やAIEージェントの運用に関する知見が不可欠であるが、その知識の蓄積・共有及び支援体制が十分に整備されていない。人的支援及びナレッジ共有を含めた総合的な利用環境の整備が必要。
- AI for Scienceにおける計算資源の需要は、研究データの整備・解析・学習等を担うバッチ型処理と、研究者の日常的な思考支援等に用いる対話型処理に大別される。これらの利用を前提とした場合、学術分野全体において大規模な計算資源が必要となることが示されており、研究者がAIを活用した研究を円滑に実施するためには、当該資源を共通基盤として整備することが必要。

第4回WGでの主なポイント（知識基盤の構築）〈委員コメント等〉

【知識基盤の構築等】

課題・期待整理

- 知識基盤の構築については、論文作成や情報処理といった機能に重点が置かれているだけでは、従来型の研究プロセスを前提とすることになるので、AI for Scienceにおいて重要となる実験や探索の加速との関係を整理することも必要である。
- データ処理の流れが直線的に整理されるだけでなく、実際のAI活用研究は、データ取得、解析、実験を繰り返す循環的なプロセスで進むものであり、こうした前提が十分に反映されることを期待する。
- AI for Scienceにおいてはロボティクスや自動実験と連携した研究の重要性が指摘されている中で、知識基盤と実験環境との関係を明確に示す必要がある。
- 分野横断的な研究の促進については重要性が認識される一方で、機能として提示されるだけでは分野融合は自律的には進まないため、研究者間の連携を実際に生み出す仕組みを検討する必要がある。
- AI利活用は、自然科学分野のみではなく、人文・社会科学分野を含む多様な研究分野への適用も増大していくことが想定される。AI利活用シナリオを検討するにあたっては、その点も考慮することが必要である。
- AIの導入による効率化については一定の効果が期待されるものの、それが必ずしも研究者の研究時間の増加や研究力の向上に直結するとは限らず、研究制度や業務構造との関係も含めた検討も必要である。
- セキュリティ及びデータ利活用の観点からは、オープンなAI基盤において機微データの利用が可能とされている点について、その安全性をどのように確保するのか、具体的な運用の在り方や企業データ等との関係も含めた整理が必要である。
- 知識基盤については、整備されるだけでは研究者に十分活用されるとは限らず、実際の利用促進やノウハウの普及、支援体制の整備が重要である。
- AI for Scienceの進展において重要となる各基盤の一体的な連携が重要。計算資源及び実験環境との関係を含めて、その構造を再整理し、研究全体を支える統合的な基盤として再設計する必要がある。

効果

- 知識基盤の構築については、提示された機能が実現すれば、研究活動において有用な支援が提供される可能性が高く、研究者の利活用が進むことが期待される。
- AIを活用した研究支援は、文献調査、データ整理、論文執筆等の負担軽減を通じて研究活動の効率化に資するものであり、研究環境の改善に寄与することが期待される。
- AIの導入により、研究テーマの探索や共同研究者の発見、分野横断的な連携の可能性が広がることから、従来の研究分野の枠を超えた新たな研究機会の創出につながる可能性がある。

政策文書における関連記載

- ・第7期科学技術・イノベーション基本計画（2026.3.27閣議決定）

第7期科学技術・イノベーション基本計画（情報基盤関連）

第1章 基本的考え方

5. 科学技術・イノベーション推進システムの刷新

我が国の科学技術・イノベーション力の飛躍的向上を図り、世界最高水準の科学技術・イノベーション政策を構築し、政府・アカデミア・産業界が一体となってサステナブルなエコシステムを実現する。これらの緊密な連携の下、基礎研究力の再興と戦略的技術領域への重点投資を大胆に進めるべく、科学技術・イノベーション推進システムを刷新する。あらゆるレベルで組織的な「縦割り」「自前主義」に陥っているマネジメント構造を、機能に着眼したレイヤー構造に転換していく。

(1) ヒト：世界標準の人材システムの構築

「研究開発マネジメント人材、技術職員やデータマネージャー、ファンドレイザーといった研究以外の重要な活動を担う高度専門人材を、点として孤立させず、組織変化に対応して役割を柔軟に変化させるためにも、組織として一括してマネジメントするとともに、これらの人材のキャリアパスを確立し、研究者と事務職員、高度専門人材が一体となって組織を動かす仕組みを構築する。」

(3) モノと情報：知と価値を創出する共用基盤の高度化

「AI時代の到来により研究システムの自動・自律化、遠隔化などに伴い、これまで以上に大量のデータが創出されることを見据え、論文、研究データ等の研究開発の過程で生まれる様々な情報の管理・活用を推進するため、**研究データ基盤システムを含む情報通信インフラの高度化も進める**。これらの情報を国・大学・国研・産業界を超えて利活用できるよう、**データの標準化や共通クラウドなど活用しやすい形での集約的な管理を実現**することで、研究から社会実装までのスピードを格段に高める。」

6. 第7期基本計画の方針

(3) 科学技術・イノベーション政策の推進に際しての留意事項

「世界各地では、研究者単独では実現できない多くのデータ収集による科学的発見や、社会課題解決への取組など、一般市民が科学研究の一翼を担う「シチズンサイエンス」や、民間・市民がボトムアップで取り組むオープンデータ・オープンソースの開発・活用が進展しており、我が国でも、多様な主体が自発的に研究活動に参画し活躍できる、新たな形での知識生産・価値創造を実現する環境整備を推進する。」

第7期科学技術・イノベーション基本計画（情報基盤関連）

第2章 知の基盤としての「科学の再興」

4.AI for Science による科学研究の革新

「A I 駆動型研究を支えるオープンサイエンスを更に推進するとともに、我が国の強みを生かした分野横断的・組織横断的な「AI for Science」の実装により、日本全体で知の生産性を向上させ研究者がより創造的な活動に専念できる環境を実現するとともに、**世界をリードする信頼性のある科学的成果を継続的に創出し、我が国の国際優位性・戦略的自律性を確保**することで、自律性と信頼性を備えた研究国家として、AI for Science 先進国としての地位確立を目指す」

(2) A I 駆動型研究を支えるデータの創出・活用基盤の整備

「より多くの研究者がA I を活用した研究環境を利用し、データの収集・解析の標準化も含めた高品質かつ大量のデータを継続的に生み出し活用できる研究システムの構築に向け、最先端の研究設備を集積するとともに、研究設備の自動・自律化、遠隔化による大規模なオートメーション/クラウドラボの形成を推進する。

全国の研究大学等において、コアファシリティを戦略的に整備し、先端的な研究設備・機器の整備・共用・高度化を推進する。大量なデータの学習を見据え、再現性・トレーサビリティ等が確保された、生物遺伝資源や機械可読な材料の実験・計測データ（メタデータを含む。）も含めた良質なデータを創出・利活用する基盤を安定的に確保・供給していく。くわえて、知の継承や海外流失の防止も含め、**電子化されていないデータやレガシーデータの利活用などについても検討**する。AI for Science の推進においては、**オープン・アンド・クローズ戦略の下で、研究データの管理・利活用を推進**する。その上で、日本の持つ研究データがA I の活用等により意図せず流出し、我が国の国際優位性などが損なわれることを防ぐため、**研究データの扱いについては、A I 技術の進展や研究分野・データの特性等に留意しつつ、研究データの国外移転、学習利用、サーバーの場所等について、国として考え方を示し、適切な管理・利活用の徹底**を図る。」

(3) AI for Science を支える次世代情報基盤の構築

「**我が国の誇る研究データの管理・利活用のための研究データ基盤（NII RDC）、流通基盤 SINET、『富岳』等のスーパーコンピュータ群を活用するとともに、HPCI（革新的ハイパフォーマンス・コンピューティング・インフラ）を中心としたユーザビリティの高い共用計算資源の戦略的な増強**を行う。「富岳」の次世代フラッグシップシステムの開発・運用を通じ、A I 処理能力・アプリケーション実行性能の飛躍と国産技術の国際市場への訴求を図る。

研究システムの自動・自律化、遠隔化などにより、これまで以上に大量のデータが創出されることを見据え、国際的なオープンサイエンスの潮流等も踏まえつつ、**AI for Science を支える研究データの管理・利活用と流通の在り方について検討し、A I 時代に適した研究データ基盤 NII RDC や流通基盤 SINET の高度化**のほか、低消費電力、高信頼、低遅延な次世代情報通信基盤の高度化を推進する。」

(5) 推進体制の構築等

「中核的拠点のネットワーク化、**各府省庁・プロジェクトを超えた研究データの共有・活用の仕組み作り**、日本の強みを生かした互恵的な国際連携や、産学連携などの取組を戦略的に推進する。」

第7期科学技術・イノベーション基本計画（情報基盤関連）

5. 研究施設・設備、研究資金等の改革

(1) 先端研究設備等の整備・共用・高度化の推進

「若手を含めた全国の研究者が挑戦できる魅力的な研究環境を実現するため、全国の研究大学等において、地域性や組織の強み・特色等も踏まえ、研究開発マネジメント人材及び技術職員を含めたコアファシリティを戦略的に整備する。研究設備・機器の管理を個人から組織に転換することで、持続的に研究基盤を維持・強化し、全国の研究者の研究設備等へのアクセスを確保する。このような研究大学等は、**SINETのセキュアで大容量のネットワークで接続することとする。これにより、先端機器群のスムーズな遠隔利用が可能となり、全国の研究者の機器へのアクセスを格段に良くする。さらに、全国の先端研究機器群から生じるデータを集約することが可能となることから、これを体系的に保存し、幅広く研究者等の利用に供する。**」

(3) 学術論文等の即時オープンアクセスの推進

「学術論文及び根拠データの即時オープンアクセスを推進する。このため、「学術論文等の即時オープンアクセスの実現に向けた基本方針」（令和6年2月16日統合イノベーション戦略推進会議決定）を踏まえ、学術プラットフォーム（グローバルな学術出版社等）に対する大学、国研等を主体とする集団交渉の体制構築を支援するとともに、**これまででは大学等ごとに個々に整備していた機関リポジトリ等の情報基盤や研究成果発信プラットフォームを、日本全体で一体的に活用できる共用研究基盤として整備・充実等を進める。**」

(4) 研究評価の見直し、研究資金制度の継続的改善等

「アカデミアにおける、学協会の活動・運営負担の在り方やヒエラルキー構造、それに伴う論文のオーサーシップの在り方といった慣習の見直しについても後押しする。」

6. 基盤的経費の確保と大学改革の一体的推進等

(2) 国際卓越研究大学制度、J-PEAKS等を通じた研究大学群の形成

「**大学共同利用機関と共同利用・共同研究拠点が中心となり、個々の大学では整備・運営が困難な最先端の大型装置や大量の学術データ、貴重な資料等を全国の研究者コミュニティに提供することで、大学の枠を超えた共同研究を促進**する。それらは、我が国独自のシステムとして研究力強化に不可欠なものであるため、共同利用・共同研究システムが担うハブとしての機能強化を図る。」

第4章 科学技術と国家安全保障との有機的連携

3. 研究セキュリティの強化等

「また、重要技術の流出防止のためには、研究者や研究機関の研究活動に係るデュー・ディリジェンスだけでなく、**取り扱う研究データ等の機微性に応じた適切なアクセス制御を含め、情報システムに係るサイバーセキュリティの強化に向けて組織的に取り組むことも重要**である。そのため、大学、大学共同利用機関及び国研について、サイバーセキュリティ対策や体制整備等に関する助言・情報共有、研修・訓練の実施、事案発生時の助言等の支援を行う。特に、特定研究開発プログラムを実施する研究機関においては、**適切なサイバーセキュリティ対策の実施が必要**であり、当該プログラムの所管府省は、適切な対応を図る。また、**特定研究開発プログラムにおける研究データ等の流出を防止**するために、大学及び国研等が「政府機関等のサイバーセキュリティ対策のための統一基準」（令49和7年6月27日サイバーセキュリティ戦略本部）の考え方にに基づき、当該プログラムの性質に応じた研究データ等の取扱いを厳守できるよう、関係府省庁が連携し手順書への必要な記載を検討する。」

**AI for Science の推進に向けた基本的な戦略方針
(2026.3.31文部科学省決定)
における関連資料およびア関連記載（各論）等抜粋**

AI for Science の推進に向けた基本的な戦略方針（政策概要）

今後5年間の集中改革期間（2026～2030年度）における国家戦略



なぜ今か（背景）

- AIが研究プロセス全体を変革
- 国際競争が急速に激化
- 今後5年間の勝負期間



日本の強み

情報基盤

SINET、NII RDC、富岳NEXT、HPCI等

研究基盤

大型先端研究施設、高品質なデータ

社会基盤

製造・計測技術、暗黙知、等

日本の課題

- ✓ AI利活用の波及・浸透
- ✓ AI高度研究人材の増加
- ✓ 共用計算資源の増強
- ✓ データの効率的活用
- ✓ 信頼できるAIの追求
- ✓ スピード感



政策的な目的

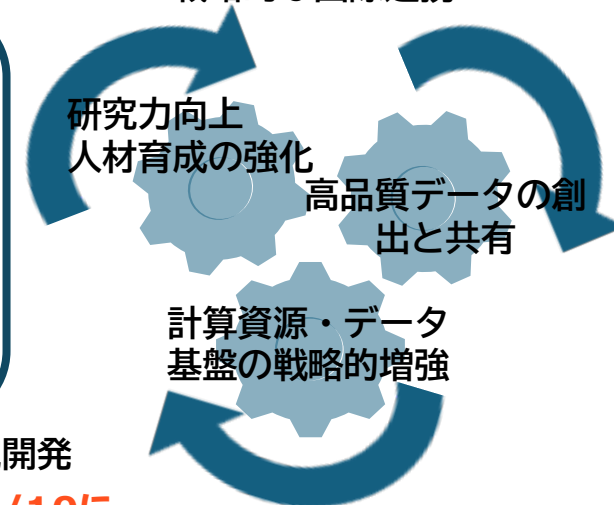
- ① 研究の質・効率の飛躍的向上
- ② 世界を先導する科学的成果の継続的創出
- ③ 国際競争力の強化・新たな価値創造

AI for Science による科学の再興

<目指す姿>

- AIが研究の自然な一部となる環境の実現
 - 分野横断的人材が学術・産業双方で活躍
 - 自律性と信頼性を備えた
- AI for Science 先進国の地位を確立

戦略的な国際連携



新たなチャレンジと普及・振興

世界を先導する研究開発

重要技術領域の先端的成果創出および研究開発期間を1/10に

将来像と期待される成果

研究プロセスの自動化・自律化、探索範囲の拡大

科学的発見の加速

複雑な現象の理解深化と新たな発見

新産業・ビジネスの創出

社会課題解決と産業競争力への貢献

国民生活の質の向上

AI for Science の推進に向けた基本的な戦略方針（概要）

- 「第7期科学技術・イノベーション基本計画」や「人工知能基本計画」、AIを巡る国際動向等を踏まえ、具体的な取組方針を整理。
- 今後5年間で集中改革期間とし、具体的な20のアクションを設定して、大胆な投資によりスピード感を持って取組を加速。
- 日本の強みを生かして、①戦略的な国際連携による世界を先導する研究開発、②新たなチャレンジとAI for Scienceの波及・振興、③これを支える次世代研究基盤の構築、④AIを高度に利活用できる研究人材の育成等を、関係省庁等と連携して強力に推進。
- 研究環境と科学研究プロセスの革新により、自律性と信頼性を備えた研究国家としてAI for Science 先進国の地位確立を目指す。

日本の強み

- **情報基盤**：世界最高水準の情報流通基盤（SINET）・研究データ基盤（NII RDC）・計算基盤（富岳・富岳NEXT・HPCI等）
- **研究基盤**：世界トップレベルの基礎科学力と多様な研究者層、世界最先端の研究装置群と大型研究施設、信頼性の高い実験・観測データの蓄積
- **社会基盤**：世界有数の経済規模、精密な製造・計測技術・ロボティクス、すり合わせや暗黙知を含む現場知、AIに対する社会的・産業的な需要

目的 I. 科学研究の革新と科学的発見の加速・質の変革 II. 研究力の抜本的強化と科学の再興 III. 国際優位性・戦略的自律性の確保

中長期的な取組目標

科学基盤モデル/エージェントやAI駆動ラボの活用により重要技術領域の先端的成果創出及び研究開発期間を1/10に

今後5年間の目標

AI for Scienceの推進により、日本の科学研究における国際優位性の確保

（ターゲット例）



3年後までに、新素材開発速度10倍の潜在力を有するAI駆動ラボシステムを開発

将来は、AI駆動ラボシステムを用いて、我が国の企業が国際的サプライチェーン上不可欠なマテリアルを量産する。



3年後までに、大規模なデータ取得を通じて、高機能なバイオ製品の高効率設計を実現するバイオ生成基盤モデルを開発

将来は、仮想細胞・生体モデルや、植物、動物、ヒト臓器等の“デジタルツインモデル”を実現し、高精度かつ高効率なバイオ製品開発や創薬等に貢献する。



3年後までに、AIIエージェント群による、最先端大型研究施設・研究装置からの大量・高品質データ創出や、仮説検証・実験の自動化・自律化を実現

新規性の高い研究を探索的に行うシステムの開発を通じて、科学研究の新しい方法論を示す。

戦略的な国際連携
（米国・英国など）

世界を先導する
科学研究成果の創出

AI for Science の波及・振興
による科学研究力の底上げ

AI for Science を支える
研究インフラの構築

（具体的な取組内容）

①研究力・人材
AI高度人材等の育成
×
AI利活用の促進

②計算資源
戦略的増強
×
利便性向上

③研究データ
高品質データの創出
×
データの一体的運用

- AI for Science のあらゆる分野での波及・振興と日本の強みを生かした重点領域の設定・投資を両輪で推進、世界トップ層との戦略的国際共同研究を推進
- AIの基礎研究含むAIそのものの研究の強化（リスク対応含む）
- 国際連携・産学連携を通じ、AI・計算資源・データに精通した人材の参画・育成、技術専門職の育成・確保、評価や処遇の見直し
- 世界最高水準の次世代AI・HPC融合プラットフォーム「富岳NEXT」の開発
- AI共用計算資源の戦略的な増強と利便性（機動性、アクセス性、相互運用性）の向上
- 産業界との連携及び国際連携を通じた計算資源の有効活用
- 戦略的価値の高いデータセットの特定・構築
- 自動化・自律化した研究設備等の整備と研究データ創出プロセスの標準化
- AI時代に即した次世代情報基盤の構築・活用、データの一体的運用

※「AI for Scienceの推進におけるAI利活用に係る研究データの取扱いに関する考え方」についても整理。

AI for Science の推進に向けた基本的な戦略方針（具体的目標例）

- 我が国の AI for Science の取組は、科学研究のあり方そのものを変える国家的挑戦。
- 第7期科学技術・イノベーション基本計画（2026～2030年度）期間となる今後5年間を集中改革期間と位置づけ、スピード感を持って推進するため具体的なアクションを設定し、大胆な投資により取組を加速する。

<研究>

① AI for Scienceの推進により世界を先導する科学研究成果を創出し、Top10%論文のうちAI関連論文数を世界3位へ（2035年度までに）

世界を先導する
科学研究成果の創出

② あらゆる分野でAI for Scienceを波及・振興し、AI関連論文数割合を世界10位→5位、AI高度研究人材を5年で3,000人増

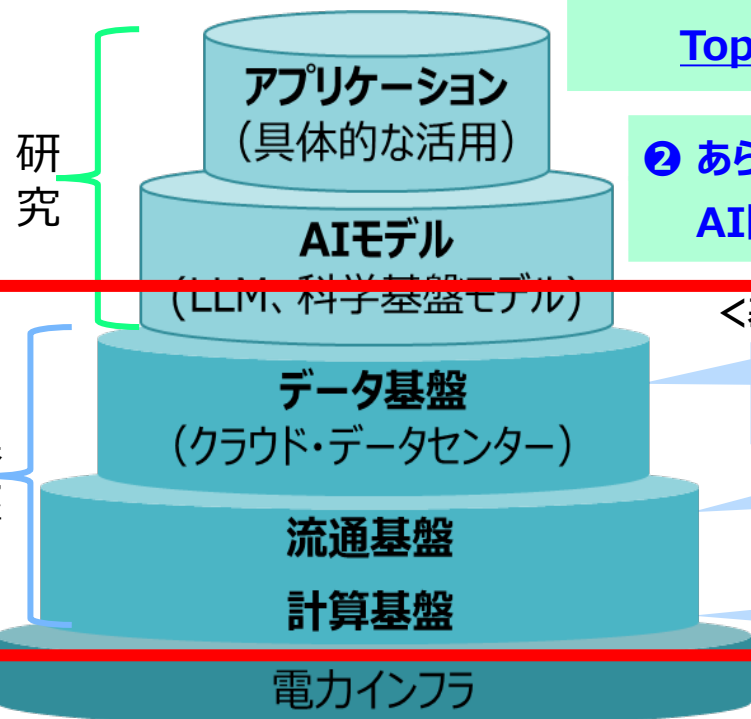
AI for Scienceの波及・振興
による科学研究力の底上げ

<基盤>

③ 研究データ基盤システムNII RDCを2030年度までに容量5倍、AI化
(※) NII Research Data Cloud

④ 学術情報ネットワークSINETを2028年度までに2倍高速化
(※) Science Information NETWORK

⑤ AI for Science 共用計算資源を2030年度までに10倍以上に



- 日本の取るべき基本戦略は、日本の資産とリソースを最大限に活用し、勝ち筋になり得る分野等の研究力を世界のトップ水準に引き上げることにある。
- そのために、国としての推進体制を構築し、研究インフラ及び研究システムを抜本的に改革する。
- あらゆる分野へAI for Scienceを波及・浸透させ、2030年には、全国どこでも誰でも、AIを駆使した高度な研究活動が可能となる社会を実現する。

AI for Science の推進に向けた基本的な戦略方針（情報基盤関連）



6. AI for Scienceの推進におけるAI利活用に係る研究データの取扱いに関する考え方

（政府方針）

人工知能基本計画（令和7年12月閣議決定）においては、「イノベーション促進とリスク対応の両立」の徹底により、「世界で最もAIを開発・活用しやすい国」を実現していくとし、また第7期科学技術・イノベーション基本計画（令和8年3月閣議決定）においては、AI駆動型研究を支えるオープンサイエンスを更に推進する旨とともに、AI for Scienceの推進においてはオープン・アンド・クローズ戦略の下で研究データの管理・利活用を推進する、としている。

（研究データの基本的考え方）

また、公的資金による研究データについては、「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）等において、オープン・アンド・クローズ戦略の下で、オープンサイエンスやFAIR原則など国際的な潮流及び慣行等に留意し、研究データを管理・利活用することを基本としており、その特性等に応じて、「公開」、「共有」、「非共有・非公開」を適切に判断することが求められている。

（AI技術の進展に伴う留意事項）

近年のAI技術の急速な進展等から、AI for Scienceの推進にあたっては、AIの利活用等を通じて、本来保護されるべき研究データや情報が予期せぬ目的に使用され、また第三者に提供される等、想定外の不利益を被る可能性があることが指摘されている。そのため、非公開とすべき研究データや、公開・非公開を慎重に検討すべき研究データに関して、AIシステム・サービスで使用する場合には、使用の可否や契約等を慎重に検討することが必要である。

（研究促進とリスク低減）

一方、研究データの取扱いについて過度な制限を設けることは、研究や国際連携の抑制につながり、結果として研究力向上やイノベーションの創出に影響を及ぼしかねない。このため、研究や国際連携を健全に推進しつつ、生じ得るリスクに応じて適切な範囲で軽減することを基本的な考え方とする。以上を踏まえ、AI for Scienceの推進におけるAI利活用に係る研究データの取扱いについて、別添2（確認項目（チェックリスト））の通り整理する。なお、研究データの取扱いについては、クラウドインフラ等の最新の技術動向や、国内外の関係諸法令及び研究動向、AIに関する議論等の動向を踏まえつつ、専門家の知見等に基づき、チェックリストを含めた本考え方について、適宜必要な見直しを行う。

AI for Science 推進に向けた基本的な戦略方針（情報基盤関連）

各論

（4）基盤：AI for Scienceを支える次世代情報基盤の構築

（計算資源の戦略的増強）

科学基盤モデルの開発・利用等の研究活動におけるA I 利活用（AI for Science）を本格的に進めるためには、A I 向けの膨大な計算資源を有する計算基盤が必要不可欠である。全国14機関が有する計算資源の共用の枠組みとして、多用な分野の研究者等が利用できるHPCIの利用状況は既に逼迫しており、計算資源の戦略的増強及び効率的な配分が喫緊の課題となっている。～ 中長期的な全体の底上げを見据えた共用計算資源の増強や、アカデミア・民間の計算資源の利活用に向けた取組を強化する。その際、**国内に整備された共用計算資源をAI for Scienceの推進に向けて最大限活用するためにも、国内で創出されたデータが高速・高信頼かつシームレスにAI学習等に取り込まれていくようなパイプラインの構築といった、データの創出基盤と計算基盤の連動についても検討**を進める。さらに、将来の計算基盤を支える先端半導体の研究開発を推進し、将来にわたっての計算基盤増強のための技術力を強化する。

（研究データ基盤の強化）

これまで以上に研究データの重要性が高まっており、**研究データをいかにA I 対応可能な形で蓄積し、利活用するかが国の研究力に直結する**。海外のA I 戦略においても、高度なA I モデルの開発等のために、A I 対応可能なデータセットの構築の必要性が明記されている。今後ますます研究データが大量に創出されることが見込まれる中、研究者自身の手でA I 等への利活用が可能となる形で保存・管理を行うのは現実的ではない。そのため、**我が国の研究データの管理・利活用の中核的なプラットフォームとして位置付けられているNII RDCについて、研究者に大きな負担をかけずにA I 対応可能な研究データを保存・管理する仕組み（AIを活用したメタデータの自動付与等）の導入等を進めていく。AI for Scienceの推進に向けてNII RDCを高度化することにより、国内外の研究機関・大学や産業界の知の結節点とし、A I 時代の国際共同研究や産学連携を促進する**。また、ライフサイエンス分野をはじめ、研究分野・対象ごとにデータの種類・形態が大きく異なることや一つ一つのデータ量が極めて大きいこと、A I 対応にはメタデータと合わせたデータ集積が必要であること等、**A I 時代に対応した一体的な研究データ基盤の整備や専門家によるキュレーション体制の確立は喫緊の課題であり、国として国際動向を踏まえつつナショナルセンター機能の強化を進める**。

（情報流通基盤の強化）

各地の**研究大学等におけるコアファシリティや大規模集積拠点における最先端の研究設備・機器から創出される高品質かつ大量のデータを、全国に張り巡らされた流通基盤を通して研究データ基盤に蓄積し、計算基盤等とのシームレスな接続によりデータの利活用を促進していくため、AI for Scienceを支える次世代研究インフラの構築は不可欠である**。さらに、**研究データ基盤NII RDCが、全国の研究者等にとって簡便に使いこなすことができる、A I 時代に適した基盤へと進化**することにより、研究データの流通は加速度的に進むため、安全・安心かつ高速に流通できる環境を実現することが必要である。こうした状況を踏まえ、**AI for Scienceを支える研究データの管理・利活用と流通の在り方について早急に検討を行い、研究者・研究機関にとっての研究活動の動脈となる流通基盤SINETの高度化を推進**する。

AI for Science の推進に向けた基本的な戦略方針（情報基盤関連）



文部科学省

（５）産学連携

産学連携では、従来の「大学が知の種を創出し、企業が実用化する」という線形型モデルを超え、科学とビジネスの近接化により、一層の加速が求められている。また、生成AIや大規模基盤モデルの登場により、科学の発展プロセスそのものが、科学的知見とAI技術が互いに補完・共創する循環型モデルへと変容している。その際、企業が必要としているのは、AIが提示する答えそのものではなく、巨額投資や開発判断を支える科学的妥当性や物理的根拠である。

このため、AI for Scienceの文脈では、大学は単なる知見提供者ではなく、AIが提示した仮説を検証・理論化する「高精度な実験・計測」と「理論構築」の共創者としての役割を担う必要がある。加えて、企業が持つ自動化・生産技術や品質管理ノウハウを大学の研究プロセスに取り入れ、AI駆動型ラボや自律型研究環境を共創することで、高速かつ持続可能な研究サイクルを確立し、産学双方で研究の生産性を引き上げることが重要である。また、イノベーションの創出のためにはインパクトのある成果が必要であり、そのためには多産多種のなアプローチも必要となる。これを加速させるためには、**様々なAI（AIモデル、AIエージェント、AI駆動ラボシステム等）が求められ、そのため、官民が一体となって計算基盤やデータ基盤等を構築していくことが必要**である。～ AI for Scienceの核心はデータの質にある。分野により状況は異なるが、非公開データや失敗事例を含む非競争領域での高品質なデータは、科学基盤モデルの構築に重要な資源となる。これを産学間で秘匿性を保ちつつ、共同活用・共有するためのプラットフォームを形成することが、新たな競争力の源泉となり得る。これにより、**産学がデータ・理論・計算資源を循環させる共創エコシステムの基盤を構築することが必要**である。

（６）国際連携・協業

日本は、ライフサイエンスやマテリアルをはじめとする様々な分野において、長年にわたり蓄積してきた高品質な研究データや、精密な計測・実験技術、世界トップクラスの計算基盤や大型研究施設等の研究インフラを有している。これらの強みを最大限に活用し、AI for Scienceの国際的な発展において不可欠な地位を確立することを目指す。～ 「AI for Scienceによる科学研究革新プログラム」を効果的に活用し、米国のジェネシス・ミッションとの連携をはじめとして、**国際共同研究、共同基盤構築、人材交流等を一体的に推進**する。

（７）推進体制

AI for Scienceの取組を強力に推進するためには、**計算資源、研究資源、人材、データ等の様々なリソースや、国際連携、産学連携、中核的拠点のネットワーク化が必要不可欠**であり、組織や分野を越えて、戦略的・統合的に推進し、科学とビジネスの好循環を作っていくことが必要である。

AI for Science の推進に向けた基本的な戦略方針（情報基盤関連）



8. 具体的なアクション

我が国のAI for Scienceの取組は、科学研究の在り方そのものを変える国家的挑戦であり、第7期科学技術・イノベーション基本計画の期間（2026～2030年度）に当たる今後5年間に「集中改革期間」と位置づけ、スピード感をもって推進するための具体的なアクションを設定し、大胆な投資により取組を加速する。

■ 計算資源

- 12 「AI for Scienceによる科学研究革新プログラム」等を通じて、研究者のニーズや用途に合わせた、産学の計算資源(GPU)も含めた機動的な利用・調達を支援する。
- 13 AI for Scienceのための共用計算資源を2030年度までに10倍以上にするとともに、HPCIについて、AI for Science時代のユーザーの利便性(機動性、アクセス性、相互運用性等)向上の取組を段階的かつ迅速に進め、2030年度までに計算資源の新たな配分システム(次世代HPCI)を構築する。この第一段階の取組として、2026年度中に、HPCIの利用制度について、AI for Scienceに関連する特定事業への採択と連動した研究資金による機動的な有償利用の実施や、申請から利用開始までの所要期間の抜本的な短縮等を図る。2027年度以降も、継続的な共用計算資源の増強や機動的な有償利用の拡大、相互運用性の向上等の取組を進める。
- 14 AI for Scienceの高度化に向けた世界最高水準の次世代AI・HPC融合プラットフォームとしての次世代フラッグシップシステム「富岳NEXT」の2030年頃までの稼働を目指す。

■ 研究データ

- 15 大学等研究機関の実験・観測等の研究活動により創出・取得される高品質な研究データについてAI-Ready化した上で、オープン・アンド・クローズ戦略を踏まえ、可能な限りFAIR原則に準拠して、利活用しやすい形で管理する。なお、競争的研究費等を通じて得られた研究データについて適切に管理・利活用されるよう大学等研究機関に周知する。
- 16 AI for Scienceの推進にとって重要な資源となる高品質なデータを大量に生成できる基盤の整備のために公募・実証プロジェクトを推進する。
- 17 開かれた研究拠点としてオートメーション/クラウドラボを5年間で少なくとも3拠点程度形成し、全国の研究者が自身のアイデアから大量のデータ生成を可能とする環境を整備する。
- 18 ネガティブデータや未整理データ等のダークデータの収集・共有・再利用に関するパイロットプログラムを実施する。
- 19 今後新たに創出される大量かつ多様な研究データについて、研究者の負担なく効率的に最大限利活用できる形で体系的に受け入れるなど、最大限利活用できるようにするため、2030年度までにNII RDCの容量を現在の5倍に増強するとともに、AI機能の付与等の高度化を行う。また、各研究分野におけるナショナルデータベースセンター等の研究基盤の整備・機能拡充を行う。
- 20 日本全国の各大学・研究機関等に存在する実験機器や計算基盤等の研究基盤を学術情報ネットワークSINETで有機的に接続し、AI for Scienceを支える最先端の教育研究環境を提供するため、AI時代に即してSINETを2028年度までに現在の2倍に高速化する。

(参考資料)

次世代の科学技術・イノベーションを支える情報基盤の在り方について (中間とりまとめ) 【概要】

背景

- 世界的に進む科学研究へのAIの応用 (AI for Science) は、産業革命と同等以上のインパクト
- 生成AIの利活用の急速な浸透により、研究DXが加速し、**将来的な研究データの流通等が質・量ともに増大**
- 社会課題解決やイノベーションの源泉である**研究データを共有し、組織・分野・セクターを超えた科学研究を行う重要性がさらに増大**

情報基盤への期待・影響

- 研究データ等の保存・管理、流通、活用を支える**情報基盤** (※) は、**AI時代の新たな科学技術・イノベーションを切り開くインフラ**となることが期待
※研究データ基盤、流通基盤、計算基盤をシームレスに接続した学術研究を支える基盤
- **AIを活用してあらゆる垣根を超えた新たな知の創造を支援し、AIが出力する情報の信頼性を担保する新たな情報基盤の構築が必要**
- 蓄積された多くの良質な研究データを学習データとして提供することで**AIモデルの更なる高度化、AI for Scienceの拡大、分野融合やすそ野の広い研究の促進**、ひいては**社会課題の解決や我が国全体の研究力・産業競争力の向上**に大きく貢献

今後の情報基盤の方向性

- 研究に伴走し、情報基盤を中心とした**研究エコシステムを支える人材**
- 組織・分野の垣根を超えた連携を具体化する**マッチング人材**
- **研究支援人材・技術者**の重要性を示し、キャリアパスとして正しく評価される**制度**



- 今後増加が見込まれるAI・データの**利用者のリテラシー向上**
- AI・データの必要性や重要性について**国全体で認識を共有**し、研究データを広く共有・活用する**インセンティブや、評価する仕組み**等の整備

AIを取り込んだエコシステムの構築

AI ⇄ データ

- **AI利用を前提**とした情報基盤へ
- 日本の文化等に理解のあるAI
- ELSI・AIガバナンスを意識した設計
- 蓄積したデータによりAIの性能を高め、AIを用いて研究し、得られたデータを情報基盤に還元してAIをさらに高度化するという**サイクルを生み出す役割**

AI for Scienceのための高度化



- 「富岳」、ポスト「富岳」、HPCI等の**計算基盤やコアファシリティ化された研究施設・設備**を情報基盤に直接接続
➡ AI for Scienceやデータ駆動型研究を加速
- **AIの普及に対応した計算基盤**や増大する研究データの流通を支える**流通基盤**



- **産業界とアカデミアの協働**の基盤となる役割
- 産学のニーズを踏まえた、**ユーザビリティを確保**した設計
- **協働が相乗効果を生む仕組み** (オープン・アンド・クローズ戦略等に留意)
- データ共有ポリシーの策定、情報セキュリティの強化、経済安全保障への対応等

効果的な配置



- **人材や認証、計算資源、ストレージ、データ管理・流通等**に関して、全国的なエコシステムとして**最適化された情報基盤の配置・整備 (集約化、分散化)** について戦略的に検討

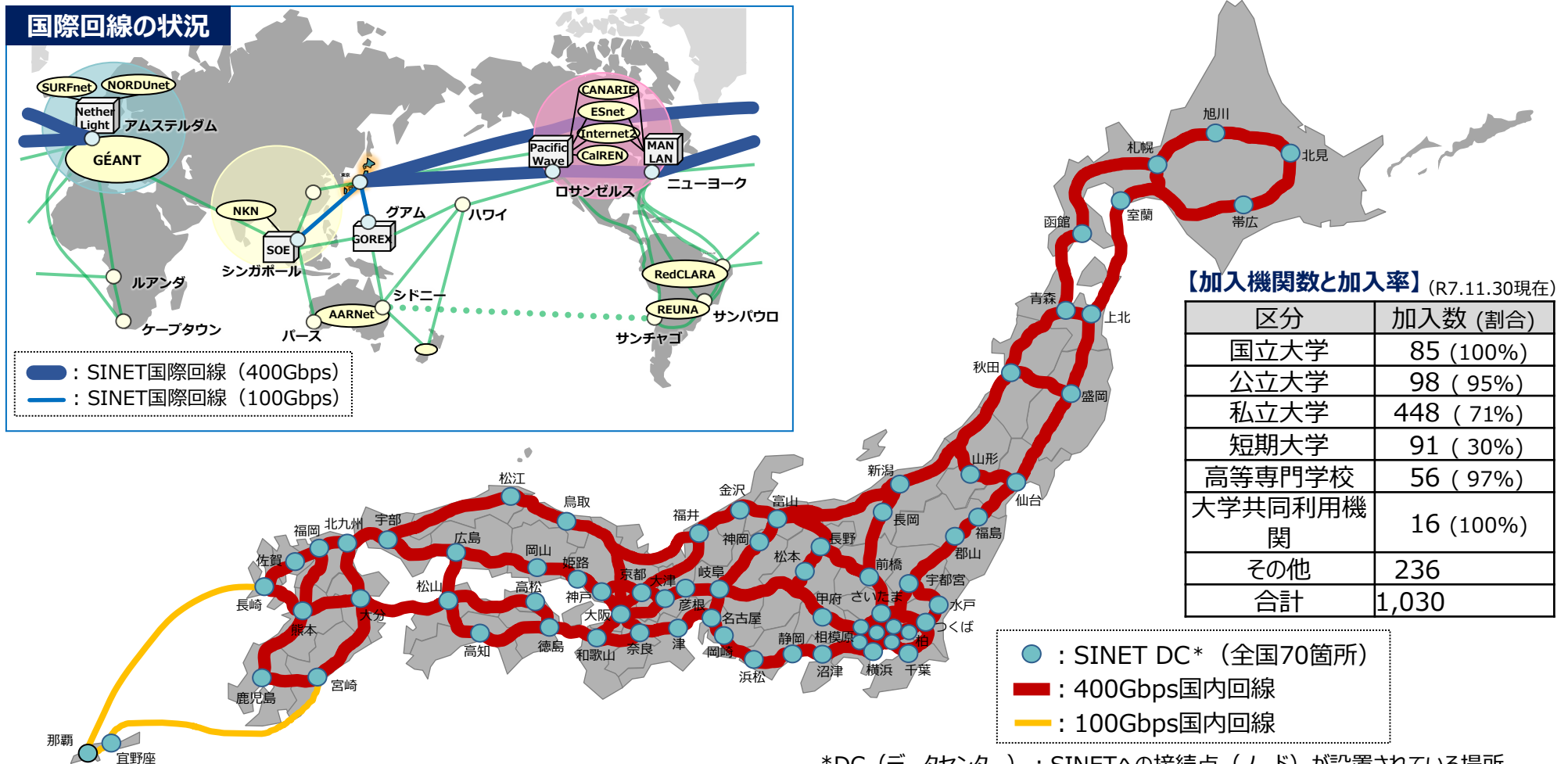
次世代情報基盤の構築を進める上でのポイント

- ✓ 長期的に**科学研究やそれを支える情報基盤のあるべき姿**を描く
- ✓ 短期的／中長期的に取り組むべき**課題と取組主体**を明確にする
- ✓ **技術進展や国際動向**に合わせて目標を**軌道修正**しながら進める
- ✓ 我が国の**独創性や潜在的な強み**を活かし、取組の強化や再構築を**速やかに進める**

学術情報ネットワーク（SINET 6）の運用

- 学術情報ネットワーク（SINET）は、国立情報学研究所（NII）が構築・運用する**情報通信ネットワーク**。日本全国の大学や研究機関等の学術情報の基盤として、**1,000以上の機関で340万人以上**が利用。大学・研究機関等との共考共創により**多様な通信サービスを開発・提供**。
- 2022年4月から、現行の**SINET6**の運用開始（日本全国を400Gbpsで接続）。2025年度より、米国国際回線は、400Gbpsへ強化。
- 研究データの管理・公開・検索を促進する研究データ基盤（NII RDC*）との融合で、データ駆動型研究ならびにオープンサイエンスの推進に貢献。今後のデータ量の増大等に対応するため、高度化・大容量化の検討が急務。

* NII RDC : NII Research Data Cloud



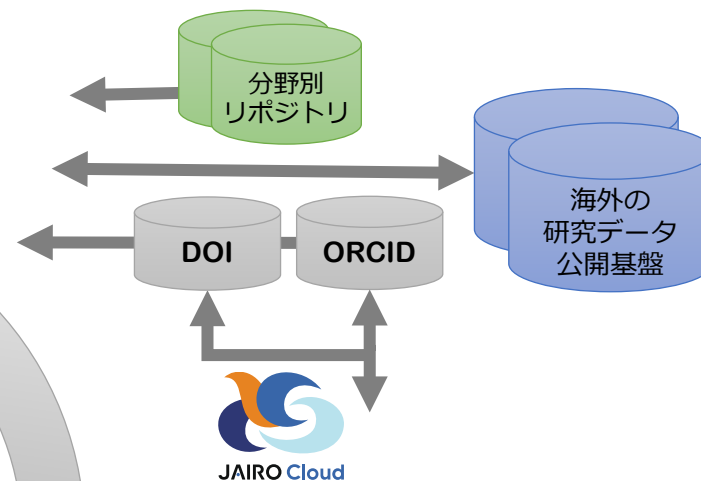
研究データ基盤の構築（NII RDC（Research Data Cloud））

開始時期：2004年（試行）～

- 機関リポジトリ等に収録された**研究論文（国内研究者論文が中心）**、**研究データや図書等を検索するためのシステム**
- 研究者や所属機関、研究プロジェクトの情報とも関連付けた知識ベースを形成
- 研究者による発見のプロセスをサポート
- 現在、年間1億3千万回以上CiNiiを用いた検索が行われている（10.7億ページビュー）（2024年）



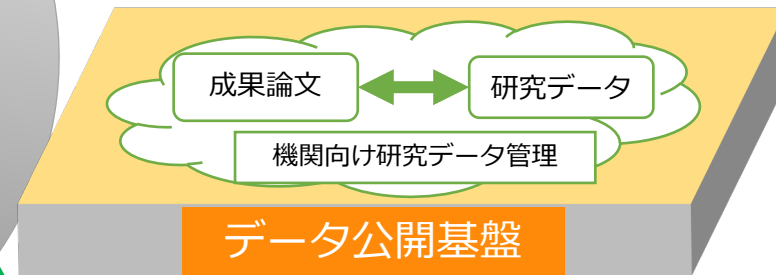
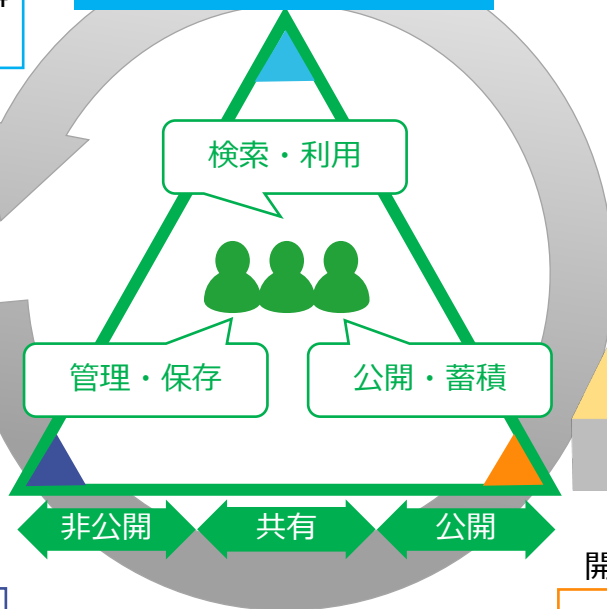
データ検索基盤



アクセスコントロール



データ管理基盤



データ公開基盤

開始時期：2019年～

- 研究遂行中の**研究データなどを共同研究者間やラボ内で共有・管理**
- 研究を進めながら適切にデータを管理することで、研究の促進や研究公正への対応を実現できる機能や、段階的な公開への準備を整えるための機能を提供
- データ収集装置や解析用計算機とも連携
- 現在、200機関が利用（2025年10月現在）

長期保存対応ストレージ領域



開始時期：2012年～

- **クラウドを使った研究成果の公開サービス**
- データ管理基盤（GakuNin RDM）との連携により、簡便な操作で研究成果の公開が可能
- NIIは大学等に、JAIRO Cloudによる機関リポジトリ構築環境を提供しており、現在808機関が利用（2025年10月現在）
- 大学等が活用することにより、研究論文や研究データの公開が促進されオープンアクセスを推進

研究データ基盤の高度化と環境整備

(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 R4-R8)

研究データ基盤高度化チーム

NII Research Data Cloudを
7つの側面から機能拡張

NII リーダ機関

研究データ基盤の機能実装

活用

コード付帯機能

研究データ基盤内で、他機関が所有する共用計算機等の解析環境を利用可能にする機能等

信頼

データプロビナンス機能

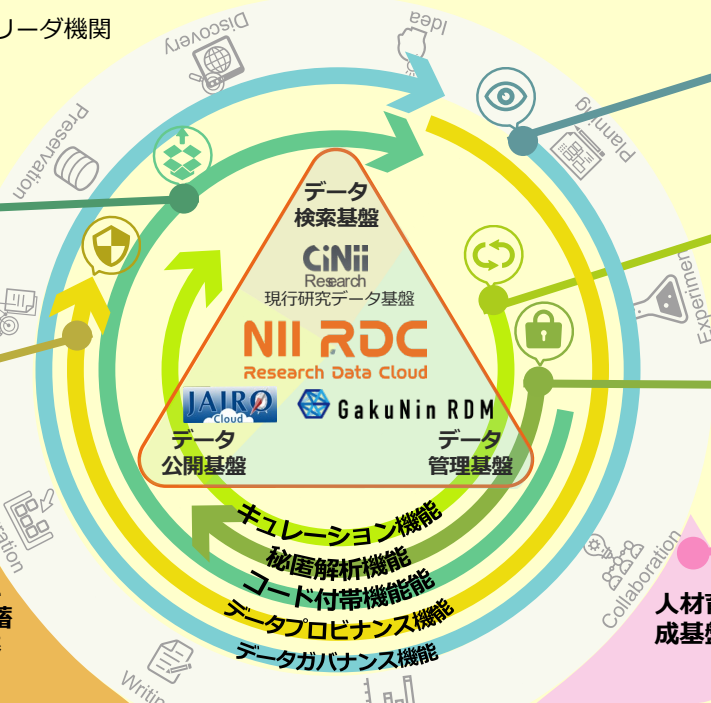
データ利用者がデータの来歴を確認できるとともに、データの提供者が自身のデータの利用状況を確認できる機能

蓄積

セキュア蓄積環境

安全で強固なデータの保存・保護機能を有する超鉄壁ストレージを提供し、機微な情報も安心して保全

セキュア蓄積環境



データガバナンス機能

管理

計画に基づきデータ管理等を機械的に支援し、データマネジメントプランをプロジェクト管理に不可欠な仕組みへと変革

キュレーション機能

流通

メタデータ付与に精通する人材と研究者を結び付け、円滑なメタデータ管理を可能とするワークフローを構築

秘匿解析機能

保護

秘密計算技術で機微な情報も安心して解析できる環境の提供で、新しいデータ駆動型研究の世界を開拓

人材育成基盤

育成

研究データ管理に必要なスキルを学ぶ環境を提供し、全ての研究者を新しい科学の実践者へと育成

プラットフォーム連携チーム



理化学研究所

リーダ機関

- ・ 機関内サービス等とNII RDCの連携機能の整理と設計
- ・ 計測機器等からの大量データを効果的に管理するための要件整理と機能開発
- ・ 管理対象となるメタデータの設計と実証
- ・ 関連する高度化機能との仕様調整と共同開発

融合・活用開拓チーム



東京大学
THE UNIVERSITY OF TOKYO

リーダ機関

- ・ 異なる分野間でのデータ活用やデータ連携に発展する取り組みを精査
- ・ 異なる分野間でのデータ活用やデータ連携に関する具体的なユースケースを創出
- ・ ユースケースをまとめたツールキットの作成とそれを用いた広報活動

ルール・ガイドライン整備チーム



名古屋大学
NAGOYA UNIVERSITY

リーダ機関

- ・ 研究データの活用に適した機械可読データの統一的な記述ルール設計
- ・ 研究データの公開に必要な要項や作業フローの整備
- ・ 研究データを適切に取扱うための指針のまとめ
- ・ 学内整備のための事例形成

人材育成チーム



大阪大学
OSAKA UNIVERSITY

リーダ機関

- ・ 人材育成を主とした研究データ管理体制の構築を押し進める学内組織構築の事例形成
- ・ 研究データ管理人材に求められる標準スキルに関する検討
- ・ 研究データ管理人材育成のためのカリキュラムの作成、オンライン学習コースの整備

基盤の活用に係る環境整備

中核機関群の代表からなる運営委員会が全体を統括し研究データエコシステムの全国展開に向けて共同実施機関を随時拡大

Before (過去)

- 網羅的なリテラチャーレビューには**限界**があり、一定の制限範囲を設けたレビューを実施した上で仮説を推敲
- 研究者の知見の範囲内**での研究計画立案（知らないやり方はできない）
- 手動での実験による、データのばらつき、時間及び人的リソースを踏まえた**限定された探索範囲**での実験の実施、再現性の問題
- 人間の知覚範囲**におけるデータ処理・分析と考察
- 上記を経た上での論文の作成例）生命・医科学分野では、**着想から論文化までの期間は約2年**



現在

- 大規模言語モデルの活用**
データ収集の範囲拡大、時間の短縮
→**情報収集の効率化**
- ロボットによる自動実験**
1つの作業を担当するロボットだけでなく、ロボット同士の連携や単独ロボットの高知能化等、AIとロボットで24時間365日実験を実行
→**探索範囲の拡大**



- スパコンによるシミュレーションデータの大量生成**
→**データ生成・分析の効率化**

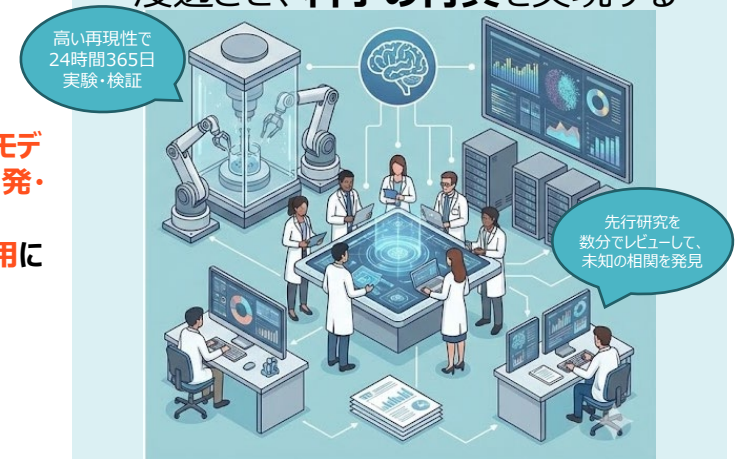


取組内容

- 【研究力】**
 - 科学研究向け**AI基盤モデル/AIエージェントの開発・高度化**
 - AIエージェント等の活用による科学研究の加速**
- 【人材】**
 - AI高度人材等の育成**
- 【計算資源】**
 - AI向けスパコン(GPU) など**計算基盤の増強**
- 【研究データ】**
 - 自動・自律実験環境等構築による**高品質データの大規模・高速創出**
 - 研究データを**AI-Ready化**

After (将来)

AI for Scienceを研究現場へ着実に浸透させ、**科学の再興**を実現する



- ✓ ユースケース**3,000件**がテンプレ化され、分野横断で再利用・横展開
- ✓ オートメーション/クラウドラボ少なくとも**3拠点程度/5年**で高品質データを大量生成、**ダークデータも資源化**
- ✓ **産学GPU・次世代HPCI・富岳NEXT**を**次期SINET**で束ね、全国で機動利用
- ✓ AI高度人材：**3,000人/5年**、GPU活用人材：**200人/5年**を育成

⇒科学の在り方の変革

- ✓ 科学研究サイクルの加速
- ✓ 論文生産性の向上と省力化
- ✓ 異分野の参入ハードルの低下
- ✓ 新たな科学的知見の創出

日本が強みを有するデータセットの例

- データの量だけではなく、中核機関に蓄積されているキュレーション等に係るノウハウや人材も強み。AI for Scienceが加速可能なのは、AI向けデータが充実している領域や、自動実験等でAI向けデータを戦略的に取得可能な領域。

■ マテリアル分野における例

NIMSデータ中核拠点（MDPF）が提供する世界最大級のデータセットの例



- 高分子材料の構造・特性を論文情報から体系的に収集したデータベース



- 物質・材料データを自動的に構造化・蓄積できるデータ基盤システム



- 無機材料の結晶構造・特性・状態図を論文情報から体系的に収録したデータベース

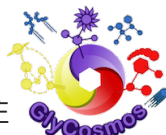


- NIMSが実施した試験により体系的に整備した金属材料の信頼性に関するデータベース

■ ゲノム、タンパク、バイオ関係（画像データ含む）における例



- 東北メディカル・メガバンク（世界初の三世代家系情報付き出生コホートを含む一般住民コホート（15万人））



- 糖鎖科学ポータル（世界初の糖鎖関連オミクスデータセット）



- ゲノム情報から、生命システム情報、疾患・医薬品情報などを統合した、京都大学が主導する、国際的にも認知度の高い、高次生命システムに関するデータベース（KEGG）



- 国際DBの一翼を担う、遺伝研のDNAデータバンク（DDBJ）

■ ロボティクス分野における例

- 一般社団法人AIロボット協会（AIRoA）がロボット動作のデータセットの公開に向けて準備中



■ 地球観測（気象・気候、防災、海洋等）等の分野における例



- 温暖化対策に向けた高解像度気候予測に関するデータベース



- 災害対応に必要とされる情報を、多様な情報源から収集したデータベース



- 海洋生物の多様性と分布情報に関するデータベース



- 全国を網羅する、陸域と海域を統合した地震・津波・火山の観測網によるデータベース



- 極域における観測や研究により創出された多種多様なデータベース

■ 最先端の大型研究施設等から創出される研究データなど



NanoTerasu



SPring-8/SACLA



J-PARC

等

■ フュージョンエネルギー分野における例

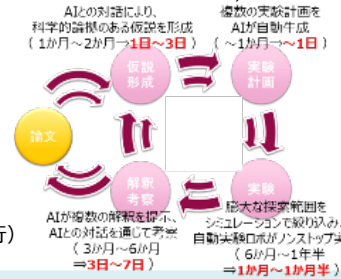
- ITER計画やBA活動への参画を通じて得られた、フュージョン分野の機器の製作や試験データ及びプラズマの挙動等に係るシミュレーションのデータ
- 世界最大のトカマク型超伝導プラズマ実験装置「JT-60SA」や臨界プラズマを達成した「JT-60」、大型ヘリカル装置(LHD)等の実験装置の建設や運転を通じて得られたデータ

「AI for Science」による科学研究の革新

令和8年度予算額 193億円
 (前年度予算額 189億円)

※運営費交付金中の推計額含む

令和7年度補正予算額 1,143億円
 ※関連経費を含めると1,527億円



現状・課題・事業目的

- 近年、AIを科学研究に組み込むことで、**研究の範囲やスピードに飛躍的向上**をもたらす「**AI for Science**」が、**創造性・効率性などの観点で科学研究の在り方に急速かつ抜本的な変革**をもたらしつつある。
- “**科学の再興**”を掲げる我が国として、AI法※の成立や急速に進展する国際潮流を踏まえ、日本固有の強みを生かした**分野横断的・組織横断的な「AI for Science」の先導的実装**に取り組むことが喫緊の課題。
- これにより、多くの意欲ある研究者及び先端的研究リソースのポテンシャルを最大化する**科学研究システムの革新**を実現し、更には産学官において広範に実装することで、我が国の**研究力・国際競争力の抜本的強化**につなげる。

事業内容：四つの柱

※人工知能関連技術の研究開発及び活用の推進に関する法律（令和7年6月一部施行、令和7年9月1日全面施行）
 ※[]内は令和7年度補正予算額

◆ AI駆動型研究開発の強化 180億円 (177億円) [490億円]

<AI基盤モデルの研究開発やデータの充実> 171億円 (169億円) [443億円]

- ライフ分野等の特定の分野に固有の強みを持つ**科学研究向けAI基盤モデル**開発や、**マテリアルデータ基盤**の充実強化等を加速。
- 科学研究向けAI基盤モデルの開発・共用 (TRIP-AGIS) 25億円 (25億円) [28億円]
- AI for Scienceを加速するマテリアル研究開発の変革 49億円 (50億円) [1億円]
- AI for Scienceのユースケース創出に向けたライフ分野の研究開発の推進 97億円 (95億円) [44億円]
- AI for Scienceによる科学研究革新プログラム [370億円]

<AI研究開発力の強化>

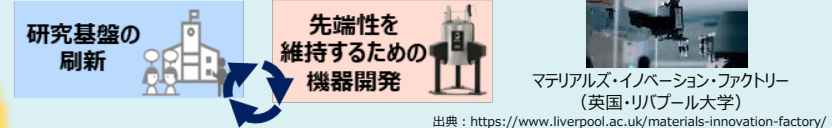
- 生成AIの**透明性・信頼性の確保**に向けた研究開発や理研AIPセンター等での**革新的なAI研究開発**を通じて「**Science for AI**」の取組を推進。
- 生成AIモデルの透明性・信頼性確保に向けた研究開発拠点形成 8億円 (8億円) [47億円]



◆ 自動・自律・遠隔化による研究データ創出・活用の高効率化 2億円[572億円]

AI駆動型研究に不可欠な高品質かつ高価値な計測データの高速かつ大規模な創出、及びその質的向上と量的拡充を図りつつ、**先端研究設備・機器の整備・共用・高度化や、大規模集積拠点の形成**を促進。

- 大規模集積研究システム形成先導プログラム 2億円[42億円] (新規) 最先端の研究設備を集積し高度かつ高効率な研究環境を実現する拠点形成により、AI時代にふさわしい研究システムの変革を先導
- 先端研究基盤刷新事業(EPOCH) [530億円] 我が国の研究基盤を刷新し、若手を含めた全国の研究者が挑戦できる魅力的な研究環境を実現するため、先端的な研究設備・機器の整備・共用・高度化を推進



◆ 「AI for Science」を支える次世代情報基盤の構築

科学研究向けAI基盤モデルの開発に不可欠な**計算基盤 (富岳NEXT・HPCIシステム等)**の開発・整備、運用や、今後大幅な増大が見込まれる**研究データの保存・管理、流通を安定的に支える研究データ基盤と流通基盤の強化**を実施。

- 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 11億円 (11億円) [5億円]
- AI for Scienceに不可欠な計算基盤の環境整備 [76億円]
- スーパーコンピュータ「富岳」及び革新的ハイパフォーマンス・コンピューティング・インフラ (HPCI) の運営及び富岳NEXTの開発・整備 177億円の内数 (181億円の内数) [385億円の内数]
- 学術情報ネットワーク (SINET) の運用 340億円の内数 (340億円の内数) [92億円の内数]



◆ 世界を先導する戦略的な産学・国際連携

AI for Scienceを世界的にリードする国内外のトップレベル機関との共同研究開発など、**戦略的な産学・国際連携体制を構築・強化**することで、**世界に伍する「AI for Science」プラットフォームの実装**を実現し、**国際プレゼンスの向上**に貢献。

- 理化学研究所における米国・アルゴンヌ国立研究所との連携 (科学研究向けAI基盤モデルの開発・共用 (TRIP-AGIS) において実施 25億円の内数 (25億円の内数) [28億円の内数])



※AI for Scienceを支える幅広い人材の育成を併せて推進。

(担当：研究振興局参事官 (情報担当) 付、科学技術・学術政策局参事官 (研究環境担当) 付、研究振興局 基礎・基盤研究課、大学研究基盤整備課、ライフサイエンス課、参事官 (ナノテクノロジー・物質・材料担当) 付)

AI for Scienceを支える次世代研究インフラの構築

研究活動におけるAI利活用(AI for Science)において「世界で最もAIの開発・活用がしやすい国」となり、「科学の再興」を果たすため、我が国の**実験基盤・データ基盤・計算基盤**を**統合的かつ戦略的に強化**するとともに、これらの基盤を**高速・高信頼・シームレスに接続し運用可能とするシステム(パイプライン)**を開発・整備することにより、**オールジャパンでの次世代研究インフラを構築**する。

【イメージ】

● 実験基盤

自動実験設備やシミュレーション、大型研究施設利用等を通じた**高品質な実験データの大量創出**

- 共用自動実験拠点を3拠点以上形成
- ダークデータの収集・共有・再利用の推進

● データ基盤

統合的な研究データ基盤の利活用による実験データの**高速なAI Ready化**や**セキュアな一元管理**

- 共用ストレージを現在の5倍以上に増強
- 各データ基盤の一体的運用と機能強化

● 計算基盤

大規模計算や高度なデータ解析を支える**世界最高レベルの利便性を備えた共用計算基盤**へのアクセス

- 共用計算資源を現在の10倍以上に増強
- システムの相互運用性や即時利用性の向上

● パイプライン

共通認証システムと**安全で高速な通信網**によるシームレスな利用、**透明性・信頼性の高いAIエージェント**等による各基盤の自動・統合運用

- 認証とUIの共通化
- データ流通の2倍高速化
- AIドリブン型基盤の開発・実装

AI for Scienceの飛躍的加速(発見・設計・検証の高速化)

産学官連携の深化によるイノベーション創出の拡大

AI時代の主導的地位の確立と国際共同研究の中核拠点化

参入障壁の大幅低下(新規研究者の増加)

ライフサイエンス分野のデータ基盤

情報・システム研究機構 国立遺伝学研究所
バイオデータ研究拠点 (BSI) ライフサイエンス統合データベース部門 (DBCLS)

五斗進・片山俊明

ライフサイエンス研究データとDBの特徴

データの多様性

- **多様な研究分野**：生物学、生化学、分子生物学、医学、薬学、生態学、農学、…
- **多様な研究対象**：ゲノム、遺伝子、タンパク質、遺伝子発現、低分子化合物、細胞、生物、環境、疾患、…
- **多様な計測装置**：DNAシーケンサー、質量分析、画像解析、…

大規模なデータ

- **大規模な計測データ**
 - 塩基配列データベース：100PB クラスのストレージが必要
 - プロテオームデータベース：1回の実験で数10TBのデータ
- **大規模な文献データ**
 - PubMed に 4,000 万件以上の論文

多種多様なDB

- **大小様々なデータベースが既に存在**

世界全体では>7000DB

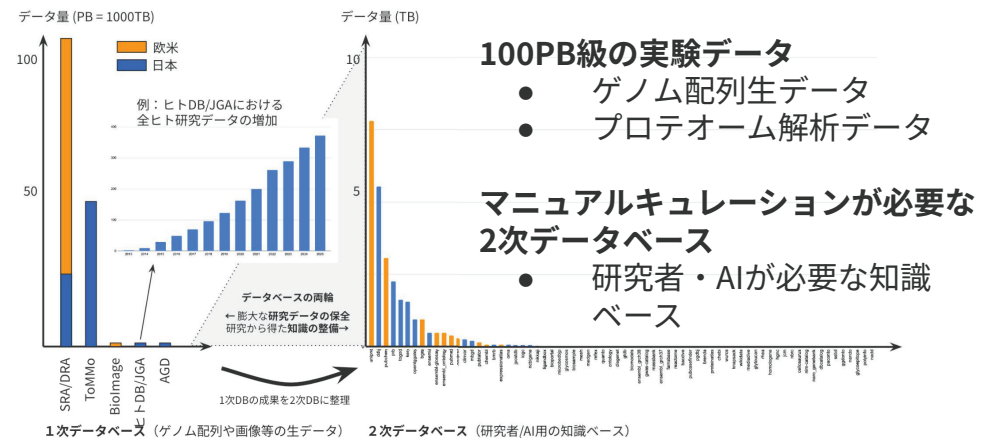
- [Integbio DBカタログ](#)
- [FairSharing](#)
- [Database Commons](#)

そのうち1329個が日本のDB
(819DBは横断検索に対応)



>1000 オントロジー

- Ontology Lookup Service
- BioPortal



ライフサイエンス統合DBプロジェクトの立ち上げと その経緯

(2001-2010 BIRD) 2006- 統合DBプロジェクト 2007-DBCCLS 2011-NBDC 2025-NLDP

目的:世界的に分散され、フォーマット、オントロジー、IFがバラバラなDBを連携・統合して
格段に使い易くすることにより、生命研究・バイオ産業の大幅な効率化を図る

・その手段として、以下を推進

- ・オープンサイエンスの推進(データの流通、データの保全、ヒトデータ共有ポリシー)
- ・統合に資する様々な技術開発
- ・カタログ、横断検索、アーカイブ、など統合の成果や各種DBのポータルサービス事業
4省連携(文科省、厚労省、経産省、農水省)の枠組み
- ・ファンディングによる統合利用を意識した多様なDB構築(本格型、育成型)
- ・国際標準化、国際連携、国内連携、技術普及活動
- ・DB人材、バイオインフォ人材の育成

・ターゲット層

- ・ライフサイエンス研究者
(基礎・応用生命研究者、医薬科学研究者、バイオインフォマティシャンなど)

FAIR原則をベースとしたデータの利活用促進



* 科学データ共有の基準としてのFAIR原則。Wilkinson, et al. (2016) Scientific Data.
DBCLSが主催している BioHackathon での議論が元になっている

Findable: データの所在を明らかにする

- Integbioデータベースカタログ (NBDC) : 約2500件のDB

Accessible: データをアクセス可能にする

- データベース横断検索 (NBDC) : 819DB
- データベースアーカイブ (NBDC) : 150DB
- 世界各地のデータベースセンター・リポジトリ
 - 米国NCBI、欧州EBI、スイスSIB、遺伝研DDBJ、京大GenomeNet/KEGG など
 - タンパク質立体構造、プロテオーム、グライコーム、メタボローム、ヒトデータベース

Interoperable: データを相互参照可能にする

Reusable: データを再利用可能にする

- データベースをダウンロードすると使える形になっている。
- データベース間で使用する用語やフォーマットを統一して使えるようにする。
- DBCLS での基盤技術開発。

統合プロジェクト、DBセンターの成果（一部）

・ライフサイエンス研究者向け AIツールなどの開発・ポータル事業

アプリケーション開発: ヒトDB(提供申請約1300, 利用申請約600 – 半数以上が海外から)、

TogoVar(20DB, 9.3億バリエーション, 25.8万人)、

PubCaseFinder(15DB, 200万データ)、TogoTV(2311動画, 300万回視聴, 1.1万人登録)、

DBカタログ(2,572DB)、DB横断検索(819DB)、アーカイブ(157DB)、など

データセット整備: TogoID(114DB, 52億IDペア)、RDFポータル(70DB, 1600億トリプル)、

PubAnnotation(1,700万文献, 650注釈プロジェクト)、TogoDX(20DB, 65属性)、など

ツール開発: SPARQL-proxy、SPARQLList、Grasp、RDF-config、TogoStanza、TogoWS、など

データセット整備やツール群開発は当初バイオインフォマティクソン用を想定していたが、その後 AI用の基盤になることが判明。現在AI利活用に不可欠の存在となった

・その他の活動

国際連携(BioHackathon15回100人, BLAH9回50人, グラフサミット6回30人, 20ヶ国)、

国内連携(バイオハッカソン16回80人, Togothon157回60人)、

トーゴーの日シンポ(500名)、AJCAS(5回, 1,700名)、など

・ファンディングにより構築された DB群

(本格型)PDBj, KEGG MEDICUS, jPOST, GlyCosmos, Shin-MassBank, SSBD, INTRARED, MicrobiomeDatahub

(育成型)ATTED-II, JoGo, DeepspaceDB, MIIB-AI, Cell IO, integMet, SSCV DB, PHI-C DB, Cura Toxii

DBサービスの一例

生命科学系データベースカタログ

- 省間連携等により収集した国内外の生命科学系DBをカタログ化
- 公開DB (日) 2,572/(英) 2,170を収載し、国内DBはほぼ網羅
- 英国Oxford大とデータ交換を実施

生命科学データベース横断検索

- 左記カタログ掲載のDBのうち、819DBを横断的に検索可能とする検索サービス

生命科学系データベースアーカイブ

- 各種プロジェクトで産出されたデータセットを公共財として維持保管するサービス
- 寄託されたDB (157件) はアーカイブ化され、研究成果が継続的に公開される

RDFポータル

- 分野横断的な研究促進に貢献するため、連携が容易で機械可読なRDF形式で統一したデータベースを集積したポータルサイトを構築
- 70件のRDF形式の生命科学データベースを用意

NBDCヒトデータベース

- ゲノム情報や画像情報等研究データを広く研究者間で共有するため、倫理面に配慮したガイドライン等を策定し、構築した国内初のプラットフォーム
- 340件の産学の研究プロジェクトからデータ提供申請
- 我が国で産出される人体由来データの収集と世界的な共有において中核的な拠点となっている

TogoVar











- さまざまなゲノムデータからバリエントを集約した、無料で自由に使えるデータベース
- 国内外のデータベースにおけるバリエントの頻度情報や、バリエントの分子生物学的アノテーション情報および既報論文をワンストップで取得できるWebサービス
- 約25万人以上のデータをもとに、約9.3億のバリエントを収録

知識グラフ(RDF)で統合された主要な生命医科学DB

共通のオントロジーやID対応関係によってAIが解釈可能なデータセット

- 塩基配列とアノテーション
 - INSDC (DDBJ/DBCLS) 
- ゲノム情報
 - Ensembl (EBI) 
 - RefSeq (TogoGenome) 
- アミノ酸配列とアノテーション
 - UniProt (SIB) 
- タンパク質立体構造
 - PDB (PDBj) 
 - BMRB (PDBj) 
 - FAMSBASE (Chuo U) 
- 化合物
 - PubChem (NCBI) 
 - ChEMBL (EBI) 
 - Nikkaji (JST) 
- 遺伝子発現
 - RefEx, GTEX (DBCLS) 
 - ExpressionAtlas (EBI) 
- サンプル
 - BioSamples (EBI/DDBJ) 
 - JCM (RIKEN) 

モデル生物
- NBRP(一部)
酵素
- BRENDA

- 医科学 ([Med2RDF](#))
 - ICGC, COSMIC, CIViC 
 - DGIdb, OpenTG-Gates 
 - ClinVar, dbSNP, dbVa 
 - ExAC, gnomAD 
 - HiNT, INstruct 
- 糖鎖
 - GlyTouCan, GlycoEpitope, WURCS, GGDonto, PAConto 
- プロテオーム
 - jPOST 
 - The Human Protein Atlas 
- パスウェイ
 - Reactome (EBI) 
- その他
 - MeSH (NCBI) 
 - BioModels (EBI) 
 - MBGD (NIBB/DBCLS) 
 - Quanto (DBCLS) 
 - :

バリエーション
- dbNSFP
- dbSNP
- TCGA

メタボローム
- MassBank

微生物・培地
- BacDive
- MediaDive
- AMR

統合DBとそれに基づく外部DBとの連携の一例

TogoVar: 日本人ゲノムバリエーション頻度の統合DB

国内の頻度DB統合

- ToMMo
- BBJ
- MGenD
- JGA

国内外の関連DB統合

- gnomAD
- ClinVar
- GWASカタログ
- MedGen
- PubMed

ハプロタイプ

- 変異の組合せ
- TogoVarで詳細

JoGo: ハプロタイプDB (九州大学)

Variant Table

実験用モデルマウス検索
● ヒトと同じ変異を持つ

MoG+: モデルマウスDB (RIKEN)

TogoID: DB間のID対応サービス

ID間の意味関係

- 114DB
- 52億ID変換ペア

海外サービス機能拡張

- TogoIDのAPI利用
- 関連DBも検索可

Id.org に変換機能提供 (欧州EBI)

汎用モジュール化

- TogoStanzaを利用

DBCLS技術提供 (欧州EBI)

PubCaseFinder: 疾患検索統合DB

疾患データ統合

- 関連遺伝子リスト

新規遺伝子パネル開発

- AMED・厚労省展開
- キュレーション

疾患オントロジー
● 電子カルテ応用
● 診断支援

難病オントロジー

- 疾患関連DB統合

NanbyoData: 難病DB
Priority-i: 重症新生児ゲノム診断 (厚労省)

マルチモーダルな生命科学DBの統合と国際情勢

海外ではナショナルセンターが統合的なデータベースの整備とサービス開発を戦略的に実施

アジアでも、インド IBDC・中国 CNCB・韓国 KOBICなどナショナルセンターの設立が進む。



ナショナル生命科学データベースセンター

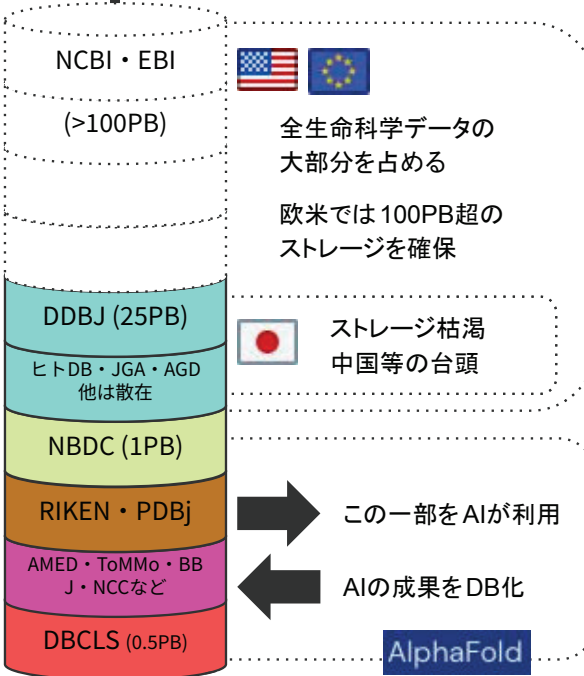


データとスパコンは直結が必須 →

ゲノム・DNAデータ

日本の生命科学データを永続的に保全する DMPと提供申請・利用申請業務

ストレージと運用を維持するための国策が必要



実験由来データとAI由来データの識別

マルチオミクスデータ


マルチモーダル・画像データ


医科学・医療データ

知識・文献データ

DBの管理機関や省庁が分散し、データが散在していること、永続的な運用が課題。予算人員に比して対応分野は広い。

欧米はナショナルセンターで国家レベルのデータマネジメントを実施、主要なデータベースが集約され相乗効果で付加価値創出や、製薬などの産業応用にも繋がっている。欧州では Elixir 支援のもと EOSC-Life とも連携して進められている。

 DDBJ+DBCLS
15億円・50人

 DDBJ, DRA, JGA
BioProject, BioSample

 AGD, CANNDs

 GlyCosmos  Microbiome Datahub

 JPOST  Mass  fanta.bio


 RDF portal  TOGO VAR  Kfoc

 TOGO ID  Kfoc

 PDBj  SSBD

 NanbyoData  dbTMM

 NCBI
300億円・350人

 GenBank, SRA, dbGaP
BioProject, BioSample

RefSeq Taxonomy

dbSNP dbVar

 GEO **Protein**


PubChem CDD


PubMed MeSH

ClinicalTrials.gov

ClinVar MedGen

GTR

 EMBL-EBI
180億円・650人

 ENA, (Federated) EGA
BioProject, BioSam

 Ensembl

 EVA  GWAS Catalog

MGNify Rfam

 ArrayExpress  PRIDE

IntAct Rhea

 MetaboLights

ChEMBL ChEBI

Europe PMC UniProt

 Identifiers.org  reactome

PDBe BioImage

 EMDB  EMPIAR

マルチモーダルな生命科学DBの統合と国際情勢

海外ではナショナルセンターが統合的なデータベースの整備とサービス開発を戦略的に実施

アジアでも、インド IBDC・中国 CNCB・韓国 KOBICなどナショナルセンターの設立が進む。

ナショナル生命科学データベースセンター



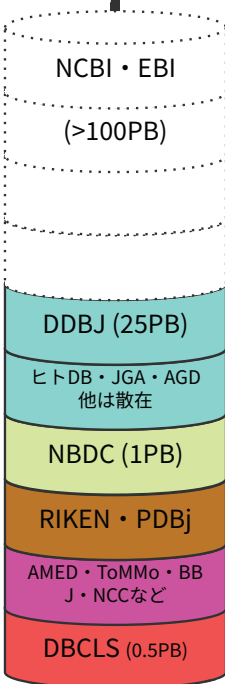
データとスパコンは直結が必須 →

塩基配列など生データのリポジトリは国際連携で開発・運用

DBの管理機関や省庁が分散し、データが散在していること、永続的な運用が課題。予算人員に比して対応分野は広い。

欧米はナショナルセンターで国家レベルのデータマネジメントを実施、主要なデータベースが集約され相乗効果で付加価値創出や、製薬などの産業応用にも繋がっている。欧州では Elixir 支援のもと EOSC-Life とも連携して進められている。

ゲノム・DNAデータ



日本の生命科学データを永続的に保全する DMP と提供申請・利用申請業務
ストレージと運用を維持するための国策が必要

全生命科学データの大部分を占める
欧米では 100PB 超のストレージを確保

ストレージ枯渇
中国等の台頭

マルチオミクスデータ

マルチモーダル・画像データ

医科学・医療データ

知識・文献データ

この一部を AI が利用

AI の成果を DB 化



実験由来データと AI 由来データの識別

DDBJ+DBCLS
15億円・50人

NCBI
300億円・350人

EMBL-EBI
180億円・650人

DDBJ, DRA, JGA
BioProject, BioSample

GenBank, SRA, dbGaP
BioProject, BioSample

ENA, (Federated) EGA
BioProject, BioSam

NII知識基盤（AI基盤モデル）との関係と期待



- **ライフサイエンス分野データ基盤と共通の基本理念**
 - オープンサイエンス
 - 研究透明性
 - データ利活用促進

- **ライフサイエンス分野データ基盤にある背景と現状**
 - データの多様性 & 大規模性
 - 専門家によるアノテーション、キュレーション、オントロジーの必要性
 - 国際連携による各研究対象に対する生データ登録用リポジトリの運用
 - 大規模データのストレージと分野に特有の解析基盤の一体的提供

- **NII知識基盤への期待**
 - データベース運用の基盤となる標準ツール（認証、トレーサビリティ、セキュリティなど）や部品の提供
 - 共通認証機構の提供：Gakunin 以上の機能。
 - 海外の研究者への提供、研究者の要件確認など
 - 分野間データ連携検索の提供：メタデータの共通化、Data Sharing Policy のトップダウン的な適用
 - 共通大規模ストレージの提供：分野ごとの事情を反映できるような環境の必要性

マテリアル分野におけるデータ基盤 蓄積から活用に向けた課題

NIMS

技術開発・共用部門長

出村雅彦

文部科学省マテリアルDXプラットフォーム事業の全体像

* マテリアル革新力強化戦略

内閣府統合イノベーション戦略推進会議決定

2021年4月27日

基本方針2：データ駆動型研究開発基盤の整備

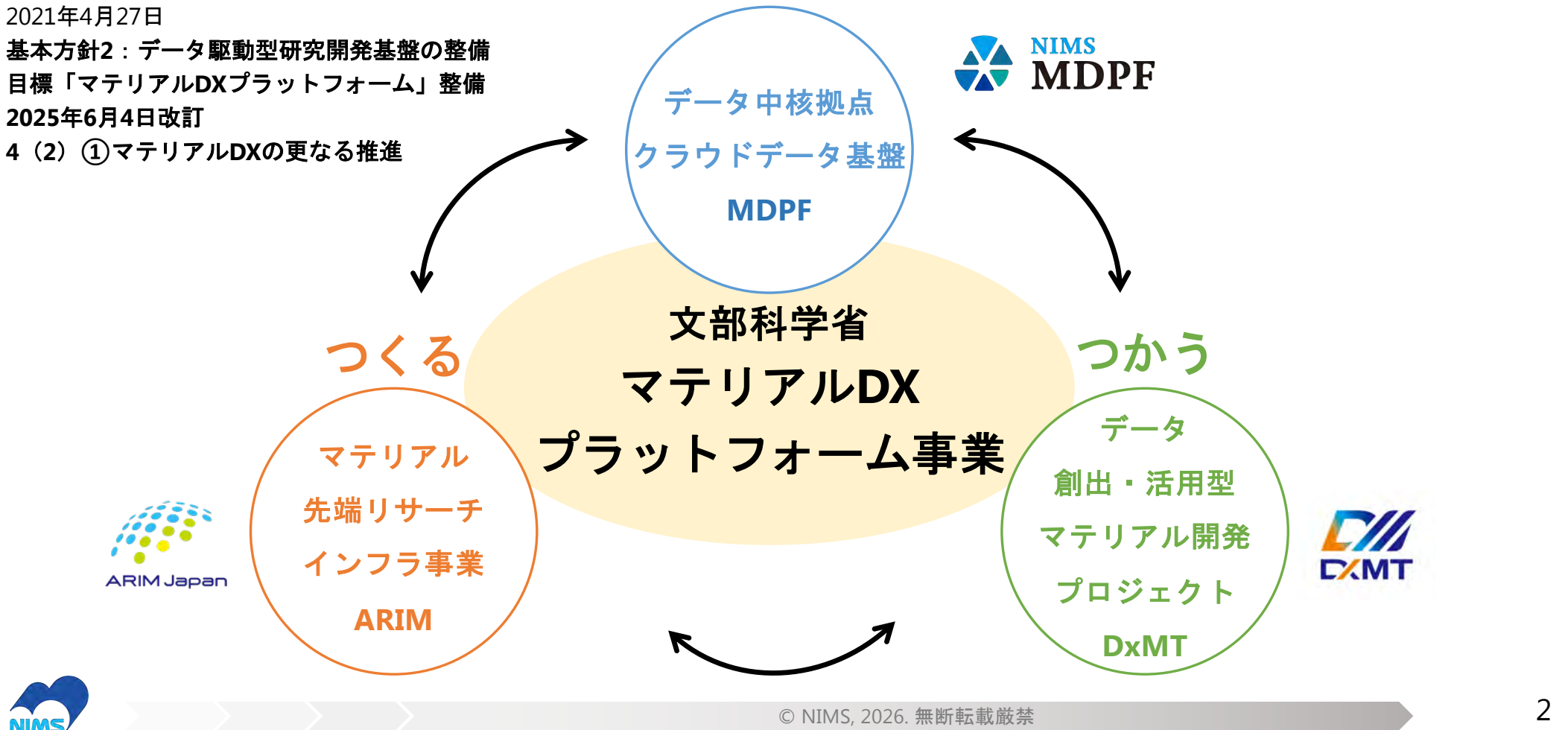
目標「マテリアルDXプラットフォーム」整備

2025年6月4日改訂

4 (2) ① マテリアルDXの更なる推進

ためる

蓄積したデータの活用ステージへ



NIMSにおけるデータ蓄積の取り組み：世界最大級の材料データベース構築

ユーザ 8,921 総アクセス 36,779,247 2023.1.17~2026.3.31

専門家がキュレーション
学術論文から

無機材料

 **世界最大**
無機材料データベース



結晶構造	状態図	特性
409,723	49,650	568,195



高分子

 **ポリマー 31,160**
物性 568,950

金属・合金

 **金属500種+機械的性質・クリープ・疲労：158,380**
 CT曲線図: 214; 硬度: 2213; 金属組織図: 627

基盤的計算データベース

- 計算状態図データベース  CPDDB
- 電子構造計算データベース  CompES-X

目的に応じたデータベースの新規開発



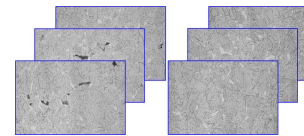
電池材料データベース
(AtomWork-Battery)

- 固体電解質 3,728
- 正極材料 1,634
- バッテリーセル 2,228



Starrydata 104,000+試料・205,000プロット

熱電材料・電池材料・磁石材料・・・



- 金属組織写真データ群



- 高温熱物性データベース
JAXA提供データ

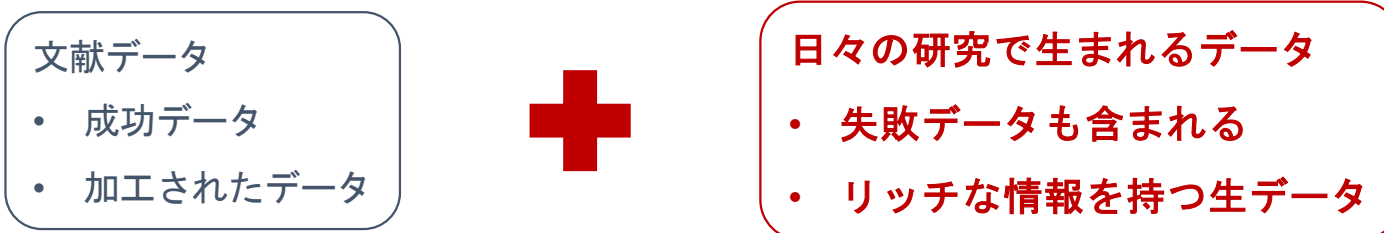
- PhononDB: 高精度フォノン計算DB 10,000+
- ホイスラ合金電子計算DB 51,000+

リファレンス
データ

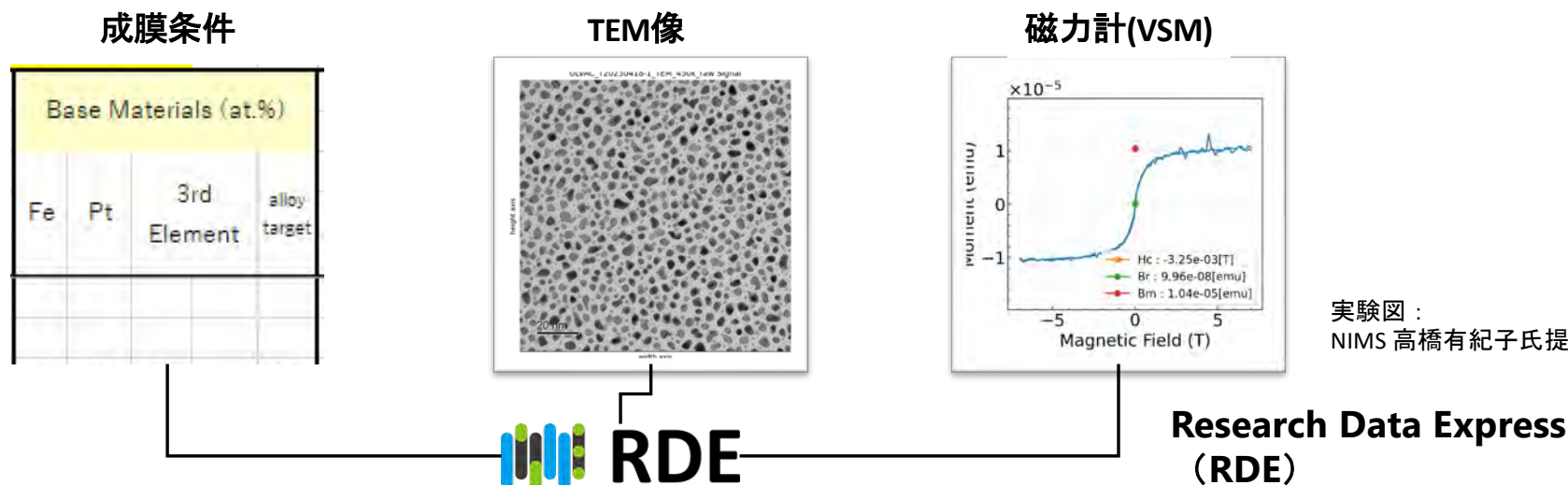


世界を先導する取り組み：文献からだけでなく、日々の研究からデータを直接蓄積

材料データをめぐる国際競争が激化：学会、出版社等 >> 研究機関の特長を活かした戦略へ



NIMSは独自にデータ構造化・収集システムを開発



RDE : マテリアルの多様な実験データを構造化する

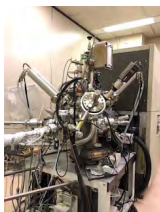
2023.1.17
リリース

装置・研究ワークフロー毎に構造化のためのテンプレートを設計



<https://dice.nims.go.jp/services/RDE/>

装置



生ファイル

測定と同時にアップロード

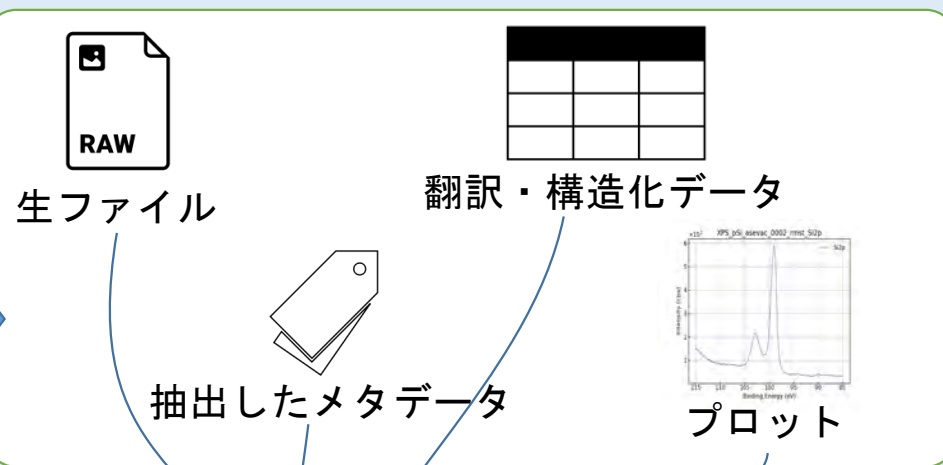
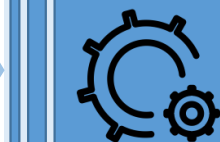


カスタムした入力フォーム



プロセス条件
実験条件
...

データセット
テンプレート

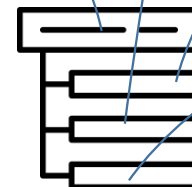


格納

アクセス制限



柔軟に共用範囲を設定可能

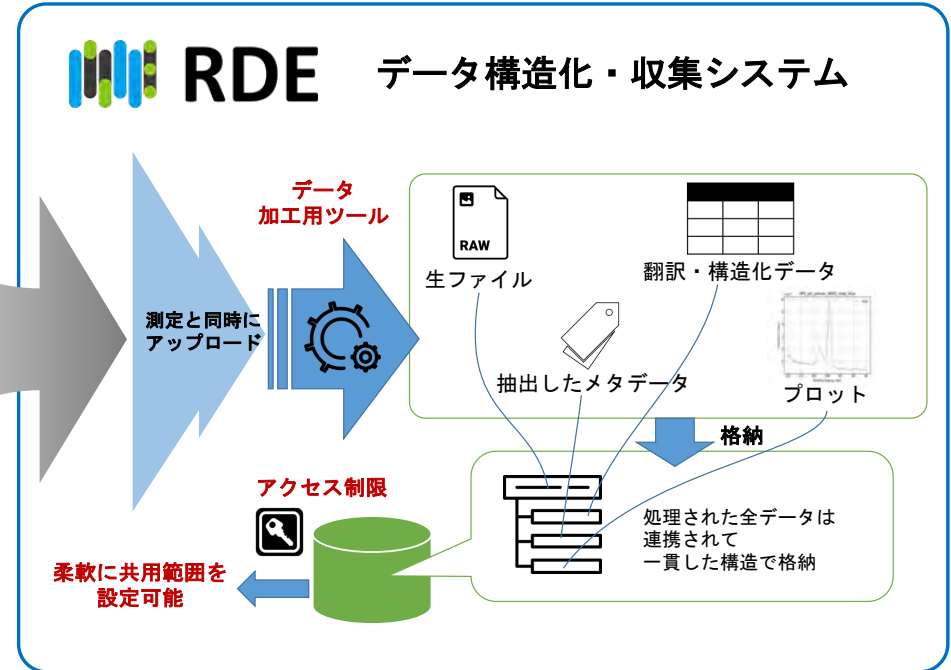
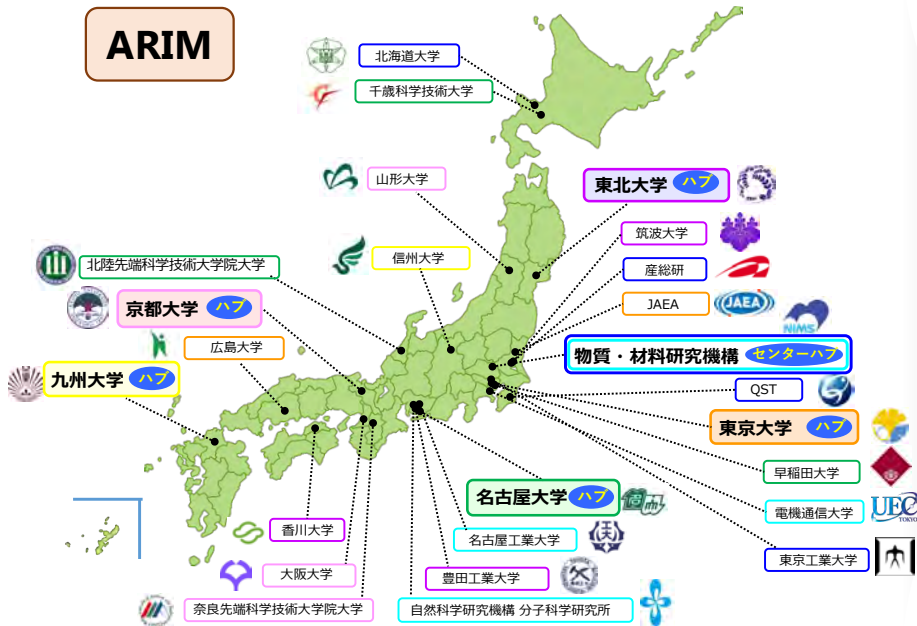


処理された全データは
連携されて
一貫した構造で格納



RDEの利用実績：研究データ登録の日本全国への広がり

データ中核拠点 



データ構造化対応装置

26機関 **1030**台

ARIM・DxMT・GteX等
国プロ事業で活用

▶ ユーザー数：5,956
データファイル：4,112,336

2026-3-31時点

RDE活用例：磁性材料の合成実験レシピと種々の計測データの集約

NIMS 高橋有紀子氏よりの資料提供

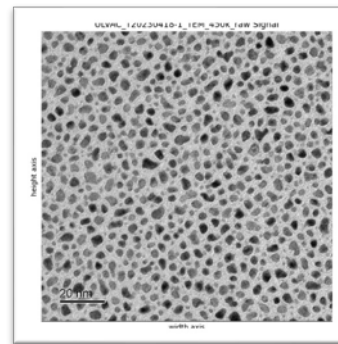
- ✓ 膜合成条件と各種評価を紐づけて管理
- ✓ 解析の自動処理で研究効率向上

成膜条件

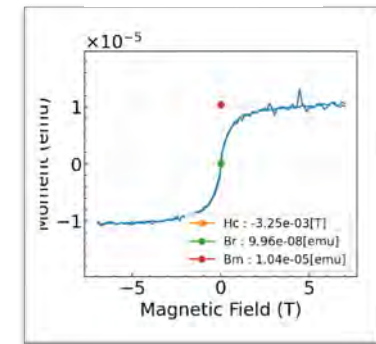
Base Materials (at.%)			
Fe	Pt	3rd Element	alloy target



TEM像



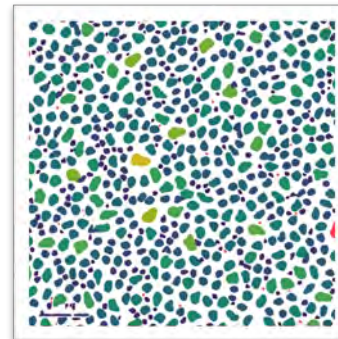
磁力計(VSM)



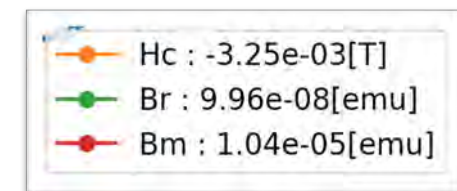
テンプレート化された電子ラボノート

登録者、サンプル名、到達真空度、基板、コメント、ガス種、ターゲット、プロセスパラメータ(レーヤー名・カソード・温度・保持時間・ガスMFC#、ガス流量、ガス圧、DF&RFパワー、成膜レート、時間、膜厚)などの記録

自動でセグメント



自動で特徴量抽出



XRD, SQUID, VSM, TEMのデータをRDEサーバーへ転送し、各データを**関連付けて蓄積**

磁化曲線を可視化し、保磁力(Hc)、残留磁化(Br)、最大磁束密度(Bm)を**自動的に**求めて記録

マテリアルDXプラットフォームが目指すデータエコシステム

エンバーゴ期間（2年）後、広域シェア



データ構造化対応装置

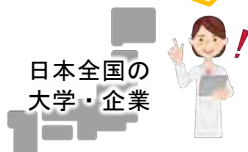
26機関 1030台

マテリアル
先端リサーチインフラ
ARIM



機械学習に向けた
形式でデータを登録

データを登録し、
様々なデータと共に
データ駆動開発!



日本全国の
大学・企業

SIP・GteX等
重要国家PJ
にて活用

計測
データ

先端
装置



データ創出・活用型
マテリアル研究開発
プロジェクト
DxMT

AI解析結果

研究データ

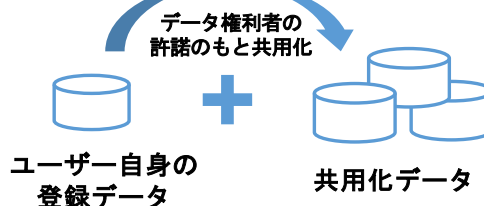
データ中核拠点 MDPF

登録ファイル数

RDE 4,112,336

収録データ数

MatNavi 2,805,703
NIMS Materials Database



世界最大級の材料データベース

2025.12一般公開

pinax クラウド上で材料データ解析と機械学習を実行



pinax

Minamoto et al. STAM: Methods **6**, 2629051 (2026). doi:10.1080/27660400.2026.2629051.

2026年3月31日時点



© NIMS, 2026. 無断転載厳禁

広域シェアサービス：世界初、論文にない実験データを提供



ARIMデータ共用サービス 2025年9月30日リリース

https://nanonet.go.jp/data_service/

世界初
リアル実験データ

AIフレンドリー
データ構造化済み

豊富なメタデータ

11万件@9/30
継続的データ拡充
+ 数十万件/年

データセット
DOI付与

多様な
マテリアルデータ
先端計測から半導体まで



お試しの
オープンデータも！

ライセンス

公共・アカデミア

一般

シングル

12,000円/年

24,000円/年

グループ
(5人以上)

60,000円/年

120,000円/年

データ利用の成果物は
利用者帰属



NIMSデータ：リファレンス・論文・ワーキングの三層構造

2026年3月31日時点

質・系統性		多様性					
リファレンス(実験)		リファレンス(計算)		論文抽出データ		研究ワーキングデータ	
Kinzoku (金属材料)	特性データ数 158,380	CompES-X (電子構造計算)	データ数 55,839	AtomWork-Adv (無機材料)	結晶構造 409,723	RDE データファイル数 4,112,336	
NIMS構造材料 データシート	258 冊	CPDDB (計算状態図) <small>*Digital-CPDDB 含</small>	データ数 1,226	AtomWork-Battery (電池材料)	材料 5,362 物質 1,117		
CCTD (CCT線図)	CCT曲線図 214	HeuslerDB (ホイスラ合金 電子計算)	データ数 51,373	Starrydata (材料物性)	試料 104,000+		
Thermophysical Property (高温熱物性)	特性データ数 385	MDR phonon calculation database (フォノン計算)	データ数 10,034	Kakusan (拡散)	データ数 31,375		
MDR XAFS DB (XAFSスペクトル)	スペクトル数 2,264	MDR lattice thermal conductivity calculation database (格子熱伝導率計)	データ数 4,811	MDR SuperCon Datasheet (超伝導材料)	レコード数 26,321		
				PoLyInfo (高分子)	ポリマー数 31,160		

無機
・
金属

高分子

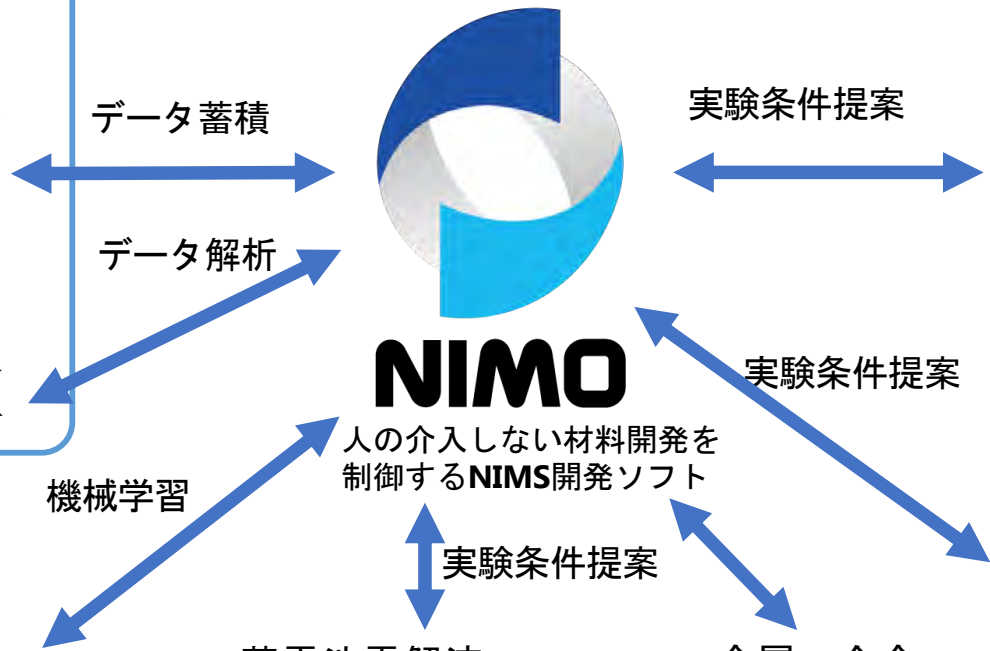


- フルオープン
- 有償
- 閲覧のみ無償/要登録 (MatNavi) :
- 広域シェア

リファレンスデータの革新：自動自律実験の拡充

中枢システム

 **MDPF**
 データ蓄積システム
 **RDE**
 AI解析基盤
 **Pinax**



磁性・半導体薄膜



ポリマー・分子合成



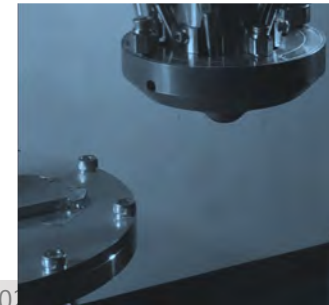
量子計算技術



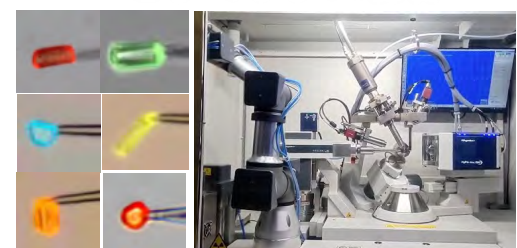
蓄電池電解液



金属・合金



セラミックス粉体



AI時代における研究データの価値変化

時代変遷	データの主用途	主な利用形態	データ価値
データ時代以前	人による参照	検索・閲覧	研究開発の補助情報
データ時代	AI解析・機械学習	データセット	AI学習資源
AI for Science時代	AIによる知識生成	AI・知識提供	戦略資源

データがAI学習資源になったことで、
バルクデータそのものが競争力を持つ時代に

NIMSでの対応例

- ・ スクレイピング制限
- ・ ライセンス提供
- ・ 広域シェア

AI時代の研究データ共有：「広域シェア」という第三の領域

全世界 オープン

- 学術基盤
 - 国際貢献
- ベンチマークデータ
→ 一部を抽出して
オープン化
 - 計算データ

広域シェア

- 認証・利用管理
 - 産学官コミュニティ共有
- 研究ワーキングデータ
→ エンバーゴ後
 - 登録・認証制
 - AI研究利用

機関・企業内 クローズド

- 競争力源泉
 - 知財保護
- 製造ノウハウ
 - 高付加価値化データ
 - 国家戦略領域

「全面公開」か「全面非公開」ではなく、
産学官共有・一部ベンチマーク化など戦略的共有が重要

多くの研究者がデータ活用の恩恵を受けるために：データ提供から知識提供へ

世界有数の材料データ基盤を整備

利用ハードルの高さ

- 専門知識・ツール操作習得コストが高い
- 探索・前処理・解析の負担が大きい

探索空間の広大さとデータの偏在

- データの蓄積は進んでいる。
- しかし材料設計空間は極めて広大
- 目的条件に直接一致するデータは限定的

次の段階としてAIによる知識活用へ

AIエージェントによる利用支援

- 自然言語でデータ検索・解析支援
- 利用プロセス自動化・簡便化

マテリアル基盤モデルによる知識活用

- 広範な材料データを横断学習した基盤モデルにより、未観測領域の特性を推定

AI支援によって「誰もが使える材料データ基盤」へ

AI for Materials構想 : Data x AI x 自動実験 = 革新的材料創出

持続可能性の確保

強靱性の確保

経済安全保障の確保

ウェルビーイング・豊かさの実現

革新的材料の創出

構造材料

超耐熱
耐水素脆化
高耐久インフラ
自己修復
AM

磁性材料

磁性熱動体
磁気記録
モーター磁石
磁気抵抗メモリ
磁気冷凍

半導体材料

二次元半導体
ニューロモーフィック
パワー半導体
有機半導体
高周波通信

ソフト材料

リサイクルポリマー
センサー・アクチュエータ
バイオアダプティブ
導電性
接着・コーティング

エネルギー材料

蓄電池材料
有機ペロブスカイト
水素製造
超伝導材料
原子力・核融合

光機能・電子材料

蛍光体材料
ダイヤモンド量子材料
光電変換
光触媒
光通信

サイバー AI-nativeデータ・ツール群



マテリアルAIプラットフォーム

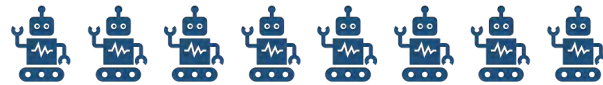
操作

データ

操作

データ

AIエージェント



マテリアル基盤モデル



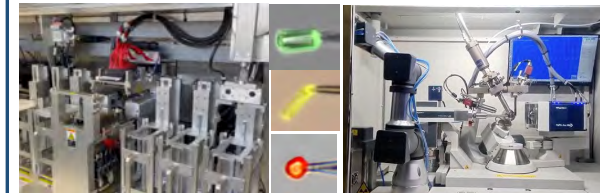
操作

データ

操作

データ

フィジカル 自動自律・先端計測・材料創製



大規模汎用基盤モデル

AI for Materials時代に向けたNIIへの期待

NIMS

マテリアル 分野

- 材料データ創出
- 材料データ蓄積
- 実験・研究現場
- データ構造化
- ドメイン知識
- 材料特化AI

学術基盤としてのNIIへの期待

分野横断 データ連携

- メタデータ連携
- 知識グラフ
- 学術知識統合

AI知識基盤 技術

- マルチモーダル
基盤モデル構築
- AIEージェント
- 知識提供基盤

データ基盤・ 信頼保証技術

- 認証・認可
- データ処理性能
- 柔軟なデータ管
理技術

求められるデータベース技術の例

これまでのRDE：設備中心＋カスタマイズ

- 共用設備：構造化を標準化
- 研究室：カスタマイズで多様な事例へ

新戦略：より柔軟に多くの研究室での活用へ

- 後から構造化で最小労力で蓄積
- Lakehouse系技術を検討

まとめ：AI時代に向けたマテリアルデータ基盤の課題と方向性

- データ蓄積から知識活用へ
 - リファレンス・論文抽出・研究ワーキングデータの三層構造で世界有数の材料データ基盤整備
 - AI時代には「データ提供」から「知識提供」への転換が重要
- フルオープンとクローズドの間として「広域シェア」・「一部ベンチマーク化」
 - AI時代には学習資源としてのデータ価値増大
 - 産学官コミュニティの中での戦略的シェア
 - 一部ベンチマーク化による戦略的オープン
- AI for Scienceに向けた次世代基盤へ
 - AIエージェント・マテリアル基盤モデル
 - AI技術・認証認可・データベース技術でNIIとの連携期待

柔軟なデータ蓄積・戦略的オープン&クローズド・AI知識化を統合した
次世代研究基盤が研究力を左右する時代へ

大型放射光施設SPring-8の研究データ基盤と AI for Scienceに向けた取り組み

初井宇記

理化学研究所・放射光科学研究センター

制御情報・データ創出基盤グループ



大型放射光施設SPring-8

SPring-8概要

- 周長**1.5 km**の加速器から高輝度**X線**を生成
- 延べ利用者数：約**1万4千/年***（約**2割**が民間利用）
- 利用分野：生命科学(薬学)、化学、材料科学、エネルギー、文化財、地球科学等
- 半導体戦略、国土強靱化、農学・食料安全保障などの国家課題解決に向けた戦略的利用を拡大
- **2029年度にSPring-8-II**へ高度化、**100倍**の高輝度化を達成見込み



データ基盤

- 分野ごとにデータ特性、オープン/クローズ方針、解析手法・ワークフローが多様
- 各分野を支える共通研究基盤が重要
- 文部科学省令和3年度補正予算により中核となるデータセンターを設置2023年9月から運用開始
広帯域ビームライン5本分（データ帯域換算で約25%）について措置
- SPring-8データセンターはデータ流通・実験中データ解析に特徴をもつ基盤
- アーカイブ、大規模データ解析はHPCIリソース、パブリッククラウドと連携
- GakuNin RDM連携サービスも試行中



データセンターを活用した成果例：

河口彰吾ら(JASRI)

大規模計算とその場測定を用いて多元セシウム塩化物を効率的に探索

課題

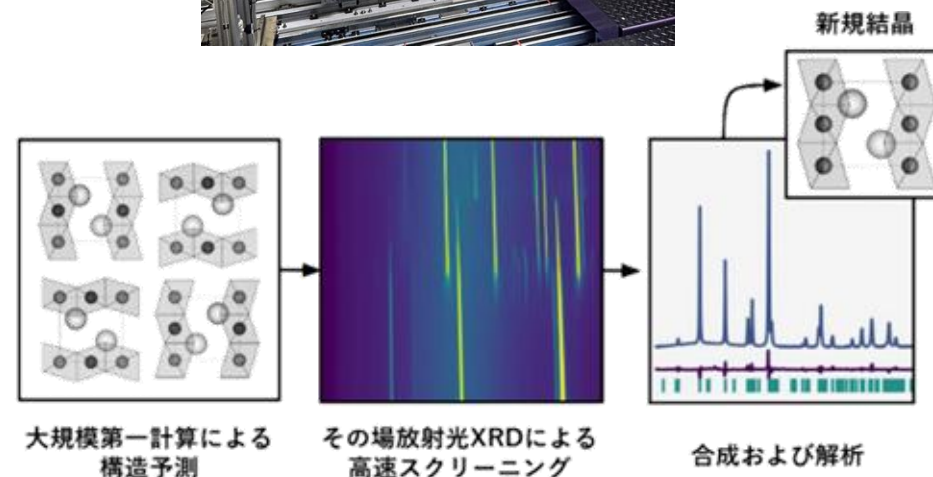
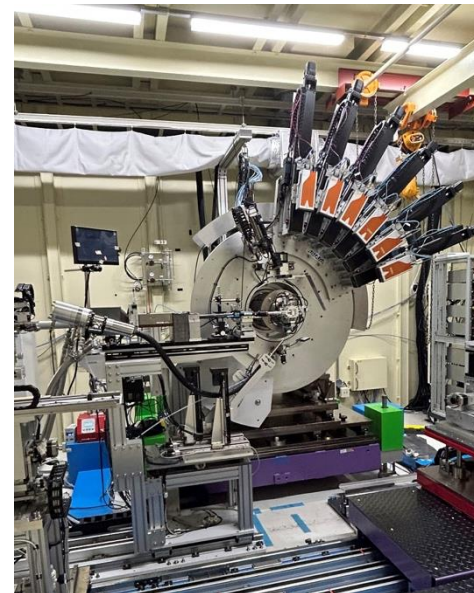
多元材料探索は組合せ爆発、実験のみでは探索効率が低い。
大規模計算と実験統合が重要

データ処理・データ量

画像データから1次元の回折プロファイルを自動処理
数百KB/データ、数千～数十万データ/日

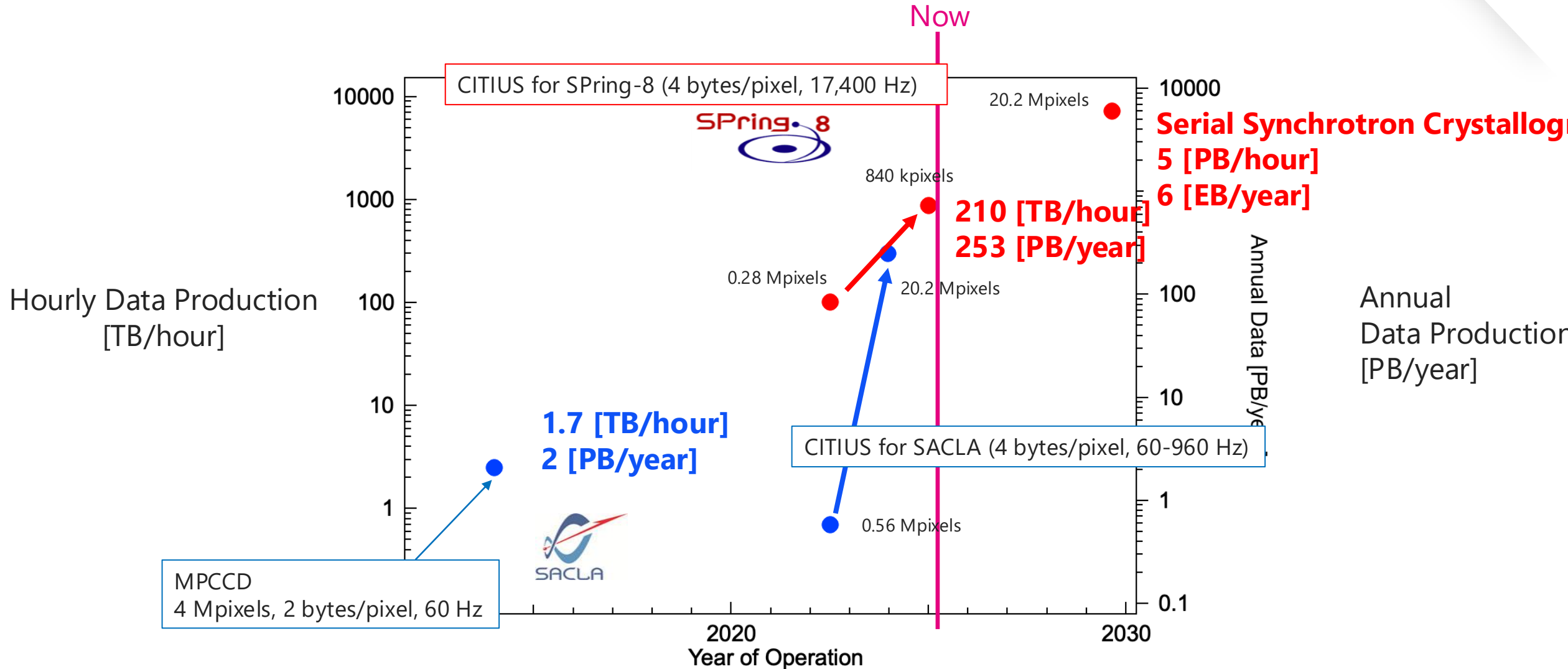
成果

第一原理計算による大規模構造予測を用いて探索
放射光X線回折による高速スクリーニング
新規セシウム塩化物の合成に成功

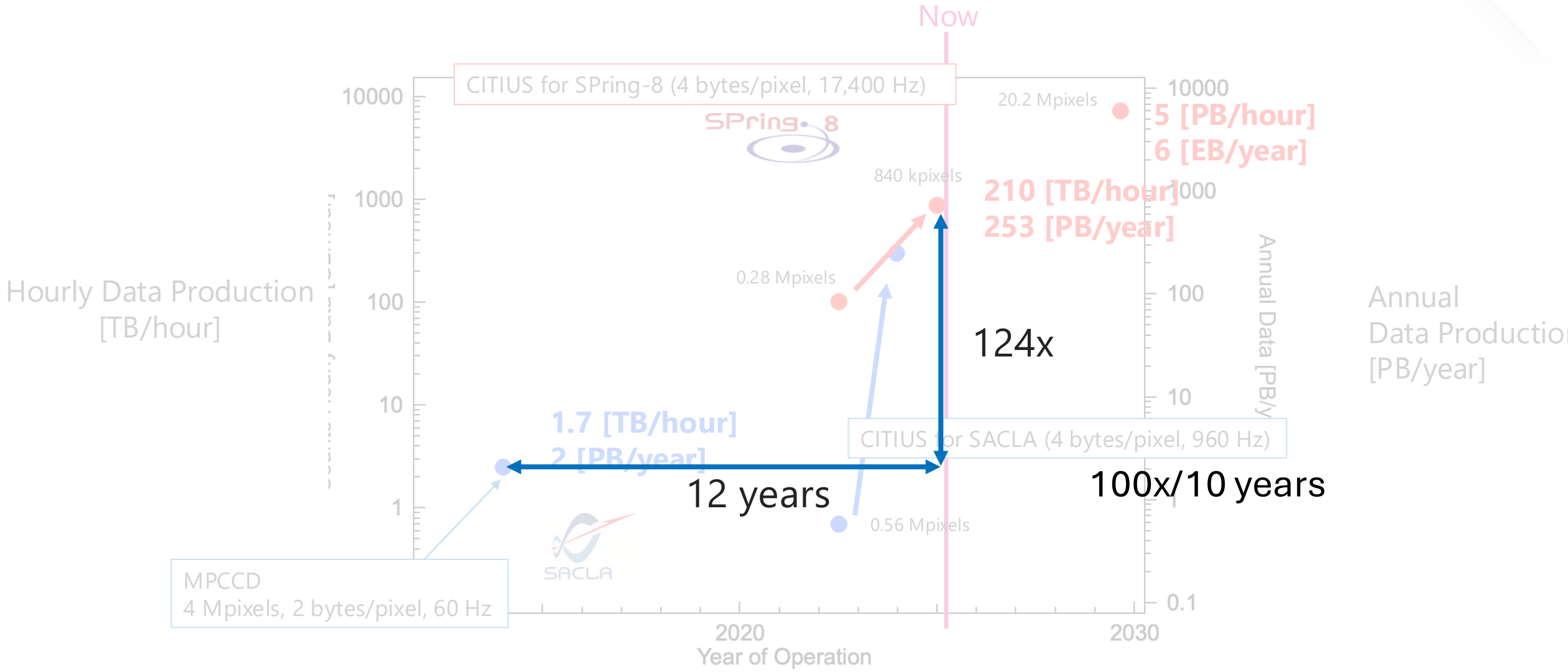


三浦章(北海道大学)、Ekin Dogus Cubuk(Google DeepMind)ら
A. Miura et al., J. Am. Chem. Soc. (2024) **146**, 29637

X線画像検出器の開発トレンド



X線画像検出器の開発トレンド



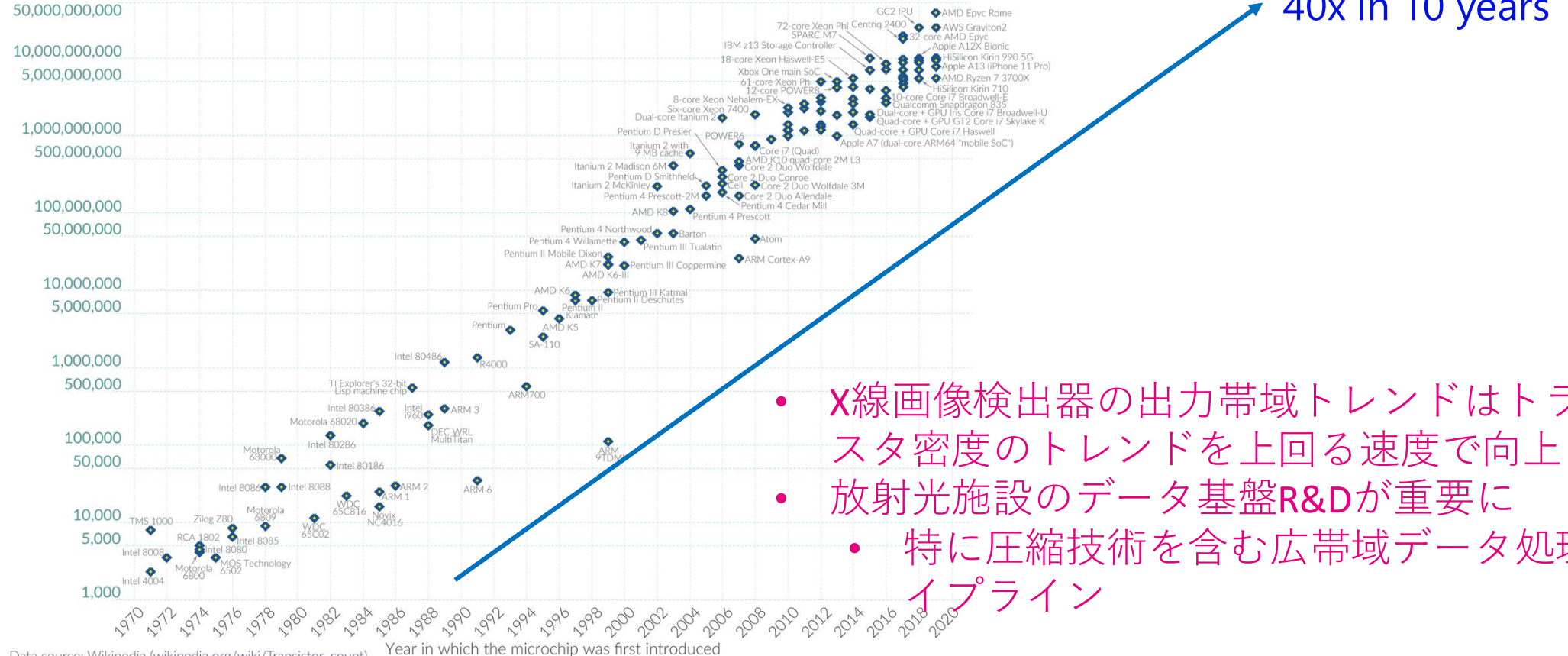
Moore's law

Transistor density 2x in 1.5-2 years

Moore's Law: The number of transistors on microchips doubles every two years **Our World in Data**

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

Transistor count



- X線画像検出器の出力帯域トレンドはトランジスタ密度のトレンドを上回る速度で向上
- 放射光施設のデータ基盤R&Dが重要に
 - 特に圧縮技術を含む広帯域データ処理パイプライン

Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)
 OurWorldinData.org – Research and data to make progress against the world's largest problems. Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

広帯域ビームライン データ処理パイプラインの実装例

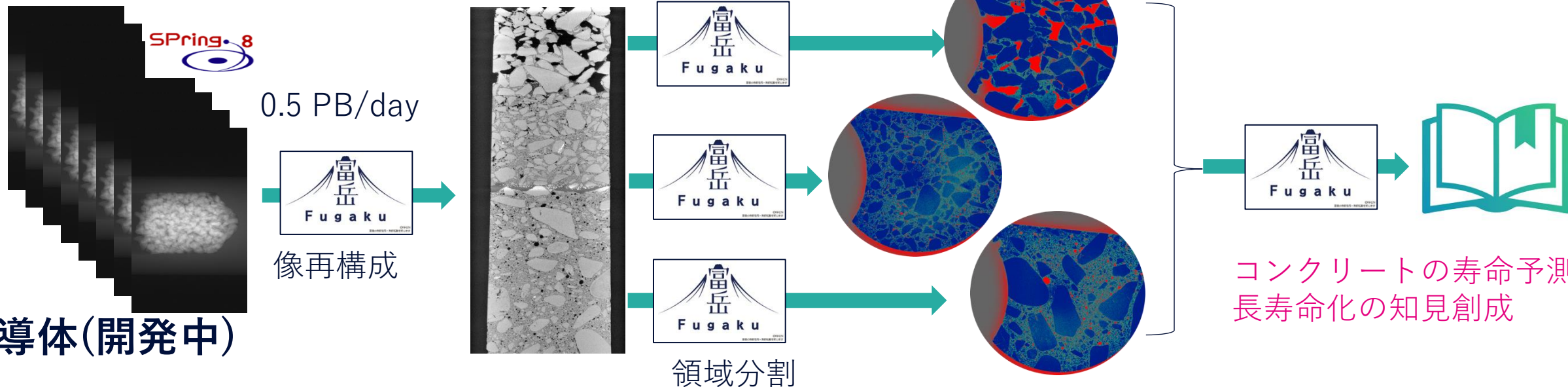
- 課題
- 理研が開発した次世代X線画像検出器CITIUSを利用した実験
 - 大量データを生成(27 GB/s, 2.3 PB/day)、リアルタイムデータ処理基盤が必須
- 現状
- 独自開発のFPGA演算加速ボード等によるリアルタイム前処理・情報抽出・データ圧縮
 - 自動データ転送とデータセンター内クラスタによる即時解析
 - ブラウザベース解析環境 (OpenOnDemand)
 - 実験条件・解析履歴を含む構造化されたデータとして保存される
- 今後
- 米国Argonne研究所APS施設に本データ処理パイプラインの技術供与・立ち上げ中
 - APS施設と共同で、AIによるリアルタイム解析・実験制御を見据えた基盤技術を開発中
 - 米国Genesisミッションで策定されるデータ形式との相互運用を検討中



国家課題解決に向けたトップダウン型の戦略的利用 SPring-8(SPring-8-II) × 富岳 (富岳Next)

コンクリート(運用中)

Mohamed Wahib (R-CCS, RIKEN)ら



半導体(開発中)

Yasumasa Joti (RIKEN, JASRI)ら



AI-Powered SPring-8に向けて：課題

データの移動

ビームライン、SPring-8データセンター、富岳と3箇所のストレージをデータが移動、ステージングがユーザの利便性を損ねている

セキュリティ

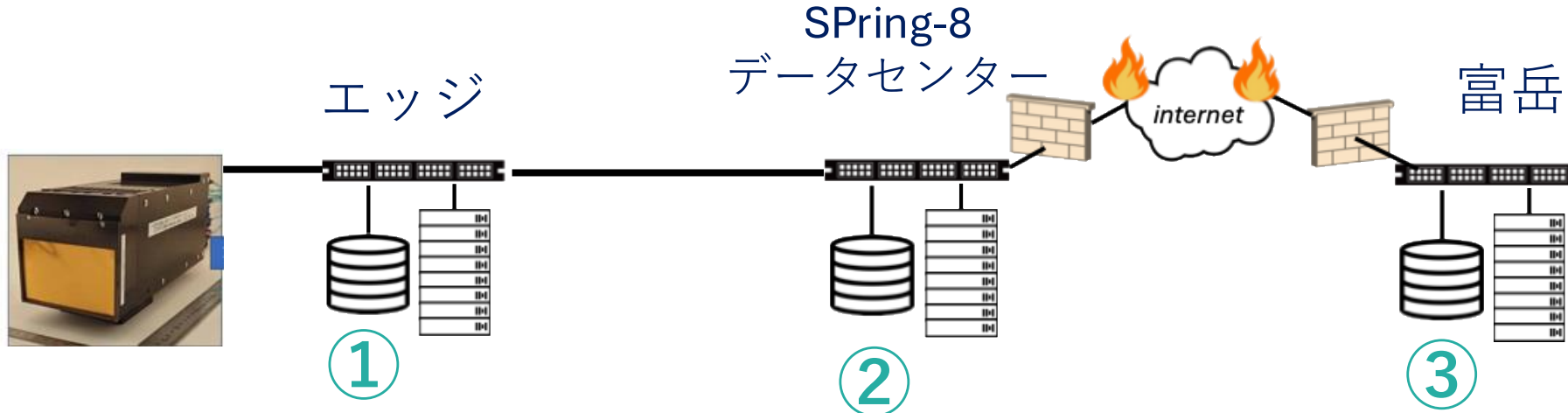
先端半導体や電池開発などの産業利用では高度なセキュリティが重要。既存インフラでは対応出来ない状況

民間クラウド事業者の**Best practice**と同等の運用・サービスが求められている
(サーバ設置室の警備員配置、監視システム、ユーザ領域の物理層での隔離オプション、監査、補償、事業継続措置等)

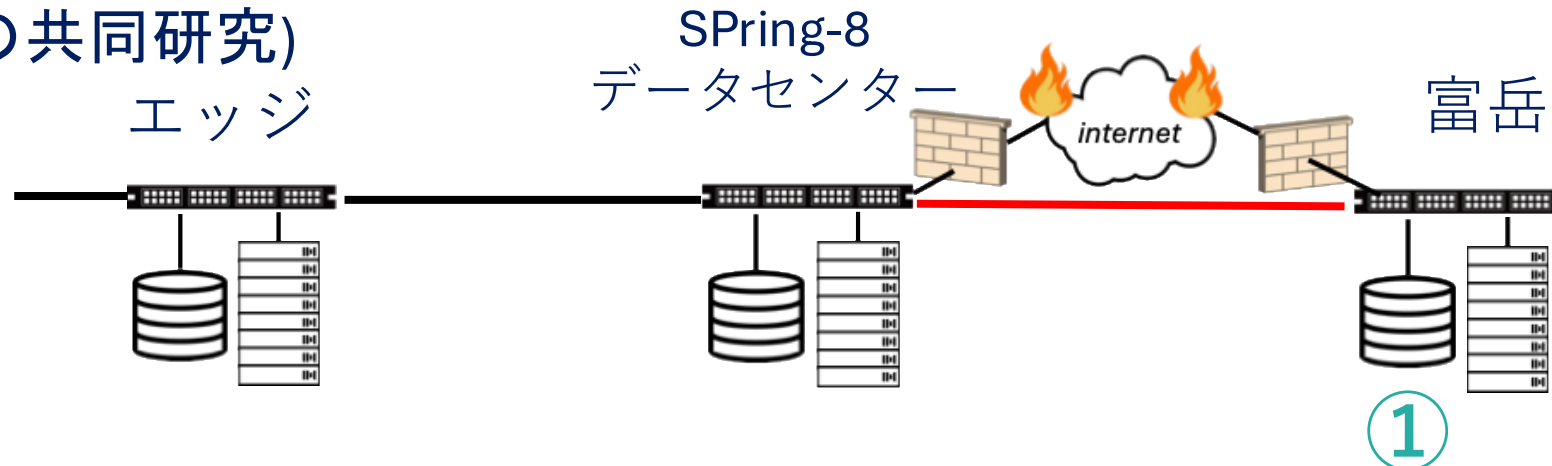
機密保持に対応出来る**LLM**が必要

セキュリティ強化： SPring-8から富岳へのデータ転送

現状



開発中(NTTとの共同研究)



特定波長を独占的に利用できるため、専用線と同様に物理的に通信経路を隔離できる。
さらに、低レイテンシ特性を活かすことで、高性能で利便性の高いストレージ構築も原理的に可能。
(ステージングの解消)

SPring-8とデータ基盤：現状

データベース

- **SPring-8**は多くの分野で利用されている。
- タンパク質結晶構造解析など、**AI**時代に有用なデータの大量創成拠点
- ユーザコミュニティが主体となってデータベースを構築、**SPring-8**は支援
データのオープン・クローズはユーザ(実験責任者)が判断できる運用

データ基盤

- **SPring-8**データセンターはデータ流通・実験中データ解析に特徴
- アーカイブ、大規模データ解析は、**HPCI**リソースおよびパブリッククラウドと連携
- データ管理について、**GakuNin RDM**連携サービスを試行中
- データ構造化への取組は重要。**AI**可読な形式でデータ保存。順次共用装置について対応中

AI for Scienceに向けた課題

利用者へのファーストタッチ

国内**9**施設に多数・多様な分析装置が分布。活用するためのコンシエルジュ機能が必須

認証

SPring-8では独自の認証基盤を運用。現在更新を検討している。民間ユーザ含む対応が必須。全施設で共通の認証基盤が必要

ネットワーク

SPring-8データセンター、富岳、**HPCI**ストレージなどが活用されている。これらを接続する**SINET**の重要性は今後増大する。特に**IOWN**の低遅延・物理層分離ネットワークによるセキュリティ強化とステージング解消等の新規機能に期待している。

AI向け計算リソース

セグメンテーション向けの**Vision Foundation Model**の暫定見積もり (**training 20M GPU/hours**、**inference 40 M GPU/hours**)。大きなリソースが必要と見積もられる。

SPARE SLIDES

AI for Science実現のための基盤整備について

2026年5月22日

国立情報学研究所

発表の目的

第1回～4回WG：

- NIIが運営する情報基盤の現状と今後の計画について

本発表：

- AI for Scienceを実現するための基盤をより広い視点での概観
- 研究分野および計算資源コミュニティからの期待、海外動向の紹介
- AI for Science実現のための基盤整備の必要性

AI for Science のための基盤

オープンサイエンスのための基盤連携

データ創出、計算、データ基盤が連携することにより、オープンサイエンスを推進

NIIが提供中のサービス

研究者, 学生,
図書館, URA,
研究コミュニティ

- ✓ 研究計画
- ✓ 実験・観測



オートメーション/クラウドラボ
先端的研究設備・機器

データ
創出

ネットワーク

SINET

学認

NII-SOCS

計算

- ✓ データ解析
- ✓ 学習



計算資源

データ



研究データ基盤

- ✓ データ管理・共有

NII RDC

研究分野
データ基盤
データ基盤

データ連携

データ
✓ 解析結果
✓ 学習結果

データ
✓ 実験・観測
✓ 調査等

AI for Scienceのための基盤連携

データ創出、計算、データ・AIの基盤が三位一体となり、エコシステムを構築することでAI for Scienceを実現

NII RDCと知識基盤機能を一体化したサービスとして新たな名称を検討中

AI対応研究データ基盤

知識基盤機能

- ✓ AI環境・サービス提供
- ✓ モデル構築支援



研究データ基盤

- ✓ データ管理・共有

NII RDC

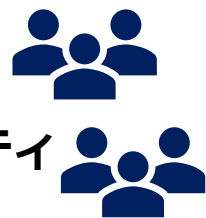
研究分野
データ基盤
データ基盤



NIIが提供中のサービス

- ✓ 研究計画・論文
- ✓ 評価・審査

研究者, 学生,
図書館, URA,
研究コミュニティ



- ✓ 実験・観測

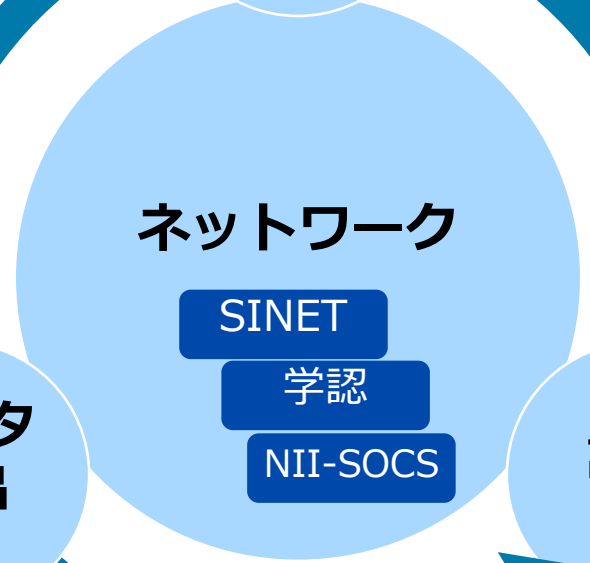


オートメーション/クラウドラボ
先端的研究設備・機器

研究活動支援

設備制御
自動化

データ
創出



ネットワーク

SINET

学認

NII-SOCS

- データ
- ✓ 実験・観測
- ✓ 調査等

計算

- データ
- ✓ 解析結果
- ✓ 学習結果

- ✓ データ解析
- ✓ 学習

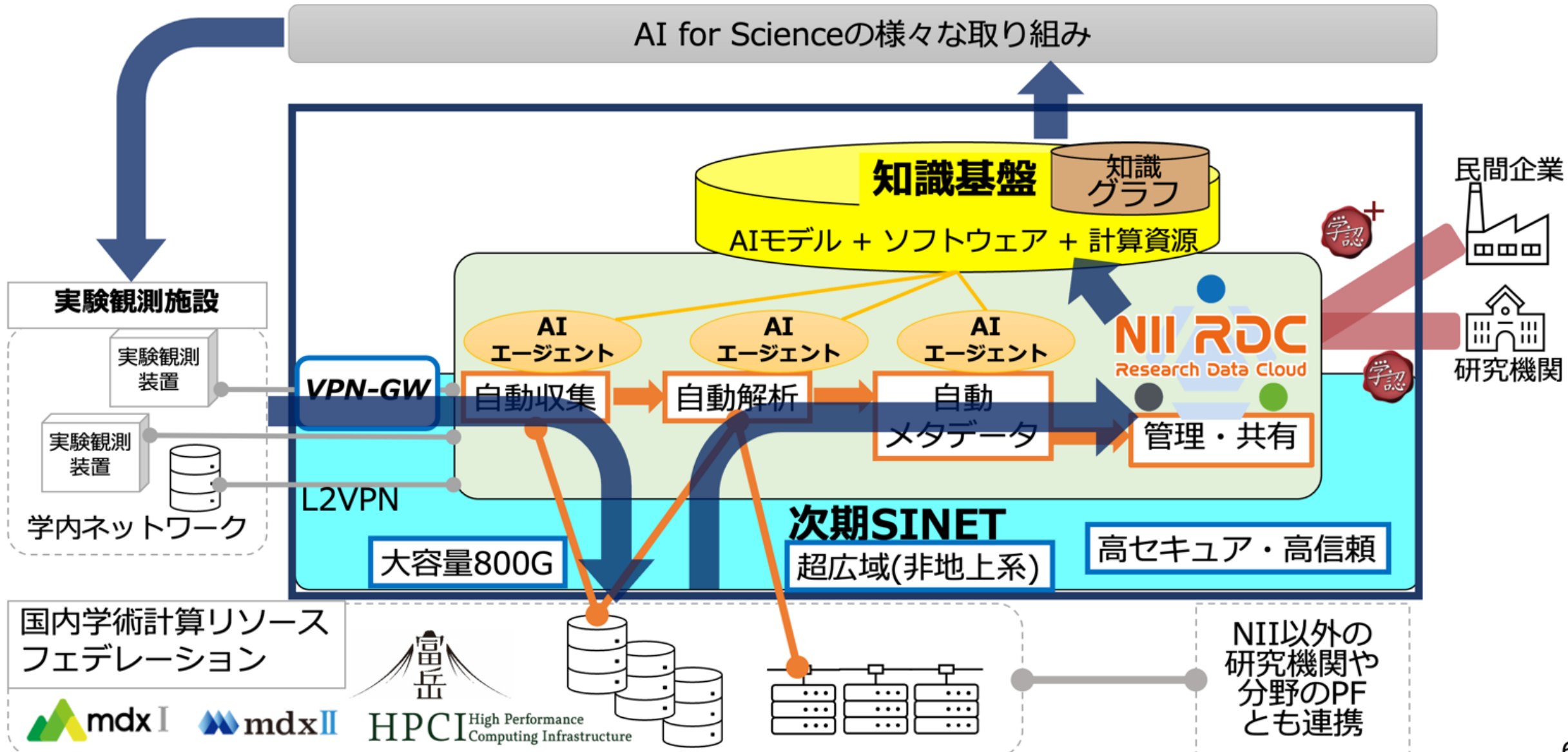


計算資源

学術研究プラットフォームの将来像

「AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ(第1回) 令和7年12月24日」資料4より改

AI for Scienceの様々な取り組み



研究分野からの期待

研究分野のデータ基盤コミュニティからのヒアリング

知識基盤に対する期待

研究分野のデータ基盤コミュニティ（ARIM、DDBJ、NanoTerasu、SPring-8、DIAS）からAIに関する取り組み、NII知識基盤への期待をヒアリングを実施

AIモデル・データ処理の高度化

- 分野特化型モデルの構築、チューニング
- メタデータ生成とデータ構造化
- 知識グラフの活用（検索・比較・発見）、結果検証

研究・実務の自動化・支援サービス

- 研究アシスト・レポート支援
- 実験自動化（エージェント機能）
- 審査業務

セキュリティとインフラ環境

- セキュリティが担保されたローカルなAI環境
- 国産モデル、OSS、モデルの信頼性
- 計算資源の確保

人材育成と知見の共有

- 専門知識の蓄積と共有
- 共同での人材育成

AIに関する現状

- マテリアル研究では、データ駆動型からAI駆動型への移行が進みつつある一方、組織全体として統一的な「AI駆動研究」体制は確立途上であり、個別研究が点在している。
- 近年は機械学習・自然言語処理に加え、自律自動AI（エージェント的手法）へと発展し、従来の解析を超えて材料候補や知見の創出（生成）に踏み込む動きが見られる。
- 基盤面では、装置・研究室ごとにばらつく実験データの構造化（AI/MLで再利用・解析しやすい形への整理）が重要課題であり、RDE等の仕組みの価値は高い。さらに、生成AI/LLMの活用により、メタデータ生成や構造化を加速する基盤モデルへの需要が顕在化している。

NII知識基盤への期待

- 大量データの整理・解析前処理・目利きには人手の限界があり、**AIによるメタデータ生成機能**の設計と、それを支える基盤モデルの整備が期待されている。具体的には、（1）**大量データを一括で整えるバッチ型のAIアシスト**、（2）**対話的に探索し「気づき」を得るサーバ型のAI活用**が想定される。NII Deep Researchには、レポート出力にとどまらず、**知識化（例：ナレッジグラフ）により多目的な検索・比較・発見**を可能にする仕組みと、利用者目線のUI設計が求められる。

AIに関する現状

- ゲノム言語モデルの構築は進めているが、GPT-Rosalindのようなパフォーマンスや機能をもつサービスの開発については未対応。
- AIを活用したメタデータの整理については段階的に実施。
- LLMCとはゲノム言語モデルなどに関する技術的な情報交換を定期的にも実施。
 - ✓ 遺伝研の分野の専門家とLLMCのAI専門家のコラボを実現。

NII知識基盤への期待

- NIIの汎用モデルを活用した、**分野モデルのチューニング**を共同で実施することに期待。
- 一般社会や企業への展開を見据えても、自然言語でやり取りできるゲノムモデルの構築が重要。
- 研究では結論の機序を明確に説明しなければならないので、**知識グラフの活用**についても期待。
- RDFだけでは情報過多になり複雑すぎて従来の検索では扱いにくいこともあるので、LLMとの組み合わせにより時代に即した発見機構を共同で構築していくことについても期待。
- ヒトゲノムなどの機密性の高い情報の扱いにおいては、**ローカルなAI環境の構築**が不可欠であり、その際のLLMCの成果の活用に期待。
- 枯渇する**計算機リソースの確保**についてもNIIに期待。

NanoTerasu

AIに関する現状

- 機械学習を使ったデータ分析の推進。
- ナノテラス・コアリションにおける研究者と企業の曖昧なニーズとマッチングのためのAIの活用。
- AIによる実験ワークフローの自動化に関する取り組み。
- AI for Industryに向けての展開。

NII知識基盤への期待

- 産業界との連携や利用を考えると、機密性の高い情報の扱いが必然。**セキュリティを担保できる環境**の提供が重要。
- カスタマイズできる、**ローカルなAI環境**の提供が必須。
- **信頼できるAIモデル**の提供も必要となり、NIIモデル利用に期待。
- AI for ScienceだけではなくてAI for Industryにも繋がる活動に繋がりたい。
- **実験と研究の自動化**を実現するためのエージェント機能の構築についても、連携した研究開発に期待。

SPring-8

AIに関する現状

- 理研のR-CCSと連携してデータ基盤を整備。ナノテラスともデータ基盤の一体化について検討。
- AIに関しては、米国アルゴンヌ国立研究所（Rick L Stevens）がけん引している状況。
 - ✓ 米国の解析ソフトウェアがオープンソースで利用可能。ガラパゴスにならないように、米国との連携を強化。
- 放射光施設として、実験的に解決する必要性がある仮説をAIに出してもらおう、Hypothesis generationが重要。
 - ✓ 膨大なデータを扱うHypothesis generationの研究開発をR-CCSと実施。

NII知識基盤への期待

- Genesis mission はDOEでの初めての大規模な民間とのコラボ。アルゴンヌとしても民間企業にはかなわないことを認識。この状況で日本はどのようにするのか？
- 企業との研究では情報漏洩に気にするので、Spring8でもオンプレで環境を提供。
- 自分たちでも動かせる**OSSなLLM**があることの重要性。**作っている人が国内**でいることも重要。
- **国内にLLM構築の専門知識を蓄積**していることが重要。あとになって、必要性を感じて人材育成から始めるのはNG。
- SPring-8での**申請課題の審査**の問題も課題。これについてもNIIのLLMの利用に期待。

AIに関する現状

- DIASのデータリポジトリに登録するデータセットのメタデータのAIによる自動付与。
- DIASのデータリポジトリのメタデータを使ったRAGの構築。チャットボットの構築。
- DIASの画像ストリーミングデータをAI解析するアプリケーションの開発など。
 - ✓ 河川の水位を画像から解析するなど

NII知識基盤への期待

- 現在のチャットボットのような機能から、データを使った総合的な研究アシストサービスの提供について検討中。そうした**サービスを構築する際のノウハウの共有**。
- 地球環境データを提供するDIASを使って、地方自治体や企業がレポートを作成する例も多々あるなど、利用者が多様。その際に、**データを理解するモデルの構築や、その出力からレポートの作成を支援するサービス**の構築について、知識基盤の機能に期待。
- DIASには、AIやLLMに強い研究者もいるので、共同して**人材育成**していく枠組みについても期待。

HPCコミュニティからの期待

HPCIコンソーシアム

- 計算機の演算性能向上に見合った容量と速度を持つ**高速ネットワーク**の確保が不可欠
- **同一のユーザID**でHPCI内の全資源にアクセスできるワンストップ・サービスを維持する
- **学術認証フェデレーション（学認）への対応**を進め、**産業界等の学術機関外IDとも連携**可能な基盤を開発する
- **大規模実験施設やIoTデータ、外部データベースと直接連携**し、リアルタイム処理を可能にする環境を目指す

The screenshot shows the website for the HPCI Consortium. At the top, there is a logo for HPCI (High Performance Computing Infrastructure) and a language selector (JA | EN). The main heading is '一般社団法人 HPCIコンソーシアム' (General Incorporated Association HPCI Consortium), with the tagline 'ユーザー視点のインフラ構築を目指して' (Aiming for infrastructure construction from the user's perspective). A navigation menu includes 'ホーム' (Home), 'コンソーシアムの概要' (About Consortium), '一般の方へ' (For General Users), '人材育成' (HR Development), 'リンク' (Link), and '本会員ページ' (Member Page). Below the navigation is a banner for 'コンソーシアムの概要' (About Consortium). The main content area is titled '提言一覧' (List of Proposals) and contains a list of proposals with their dates and titles:

- 2025年6月5日
【提言】次世代計算基盤のユーザビリティに関する提言
- 2024年6月25日
【提言】「次世代計算基盤を利用した成果の最大化に向けて」
- 2023年5月19日
【中間報告】「HPCI システムの今後の在り方」に関する調査検討について
- 2022年6月22日
【提言】「富岳」本格運用時のHPCIおよび次期フラッグシップ計算機の在り方について
- 2021年6月25日
【提言】「富岳」本格運用期における計算科学技術振興の在り方について
- 2020年6月15日
【提言】今後のHPCIシステムの構築とその利用に関する基本的な考え方について

HPCI計画推進委員会

- **SINET**を活用し、日本全国の大学や研究機関を高速回線で結ぶことで、場所を問わず高度な計算環境へのアクセスを実現
- 全国の利用者が**一つのユーザーアカウント**で、用途に応じて多様な計算資源をシームレスに利用できる環境を維持・発展
- **計算基盤、ネットワーク、データ基盤が一体**となり、研究活動のデジタルトランスフォーメーション（研究DX）を支える
- **遠隔の研究機器からのデータ取り込みや、複数拠点間での効率的なデータ連携**など、次世代型の運用ニーズへの対応

The screenshot shows the official website of the HPCI Plan Promotion Committee. The page is in English and features the following content:

- Header:** 文部科学省 (Ministry of Education, Culture, Sports, Science and Technology - Japan) with a logo and a menu icon.
- Language:** A button to switch to English.
- Search:** A search bar with a "検索" (Search) button.
- Breadcrumbs:** トップ > 政策・審議会 > 審議会情報 > 調査研究協力者会議等 (研究振興) > HPCI計画推進委員会 > 次世代計算基盤に関する報告書 最終取りまとめについて
- Main Content:**
 - **次世代計算基盤に関する報告書 最終取りまとめについて**
 - [次世代計算基盤に関する報告書 最終取りまとめ 本文 \(PDF:764KB\)](#)
 - [【別紙1】次世代計算基盤検討部会 中間とりまとめ \(PDF:426KB\)](#)
 - [【別紙2】次世代計算基盤 開発/運用に向けて 理化学研究所の果たすべき役割 \(PDF:7.6MB\)](#)
 - [次世代計算基盤に関する報告書 最終取りまとめ ポイント資料 \(PDF:450KB\)](#)
- Footer:**
 - Get Adobe Acrobat Reader button and text: PDF形式のファイルを御覧いただく場合には、Adobe Acrobat Readerが必要な場合があります。Adobe Acrobat Readerは開発元のWebページにて、無償でダウンロード可能です。
 - Navigation links: [文部科学省ホームページトップへ](#) and [ページの先頭に戻る](#)
 - Footer menu: [サイトマップ](#) | [災害関連情報](#) | [官公庁等リンク集](#) | [English](#) | [キーワード](#)
 - Footer text: 御意見・お問合せ | [プライバシーポリシー](#) | [リンク・著作権について](#) | [アクセシビリティへの対応について](#)
 - Footer text: 文部科学省 〒100-8959 東京都千代田区霞が関三丁目2番2号 電話番号：03-5253-4111(代表) 050-3772-4111 (IP 電話代表) 法人番号 7000012060001 [案内図](#)
 - Footer text: Copyright (C) Ministry of Education, Culture, Sports, Science and Technology

https://www.mext.go.jp/b_menu/shingi/chousa/shinkou/020/shiryo/mext_00003.html

海外公的機関におけるAI for Scienceを見据えた 研究インフラ（プラットフォーム）の動向

「科学技術・学術審議会 情報委員会（第47回） 令和8年5月12日 資料4」より再掲

概要

我が国のAI for Scienceを促進するための新たな情報基盤整備の参考とするため、海外の公的機関等におけるAI及び学術知識グラフを活用した学術基盤プラットフォーム構築事例に関する委託調査を以下の通り実施した。

■実施期間

- 2026年1月～3月

■調査項目

- 我が国での構築の参考とするため、学術基盤プラットフォームが、どのような政策を背景として、どのような機能・実装となっているか、AI及び学術知識グラフを活用しているかを把握する。

■調査対象

- 主要国の代表的なプラットフォーム10機関について、AI及び学術知識グラフの活用状況、公開情報の入手可能性、NII関係者のヒアリングも経て決定した。

■調査方法

- Web調査を主体とし、DataON/ScienceONについてはインタビューも実施して情報を補完した。

今回、上記調査のうちのEOSC（EU）、SURF（オランダ）、NSF ACCESS（米国）、DataON（韓国）についてNIIが編集・再構成した内容、及びNIIが追加で調査したThe American Science Cloud（米国）を報告する。

調査結果概要 (1/2)

	EOSC (EU)	AI4EOSC (EU)	SURF (オランダ)
特徴	欧州の研究者がFAIR原則に基づき、分野や国境を越えて研究データを分散型で保存・共有・再利用できる、オープンな統合プラットフォーム	EOSC上でAI/ML/DL技術を活用するためのサービスを拡充し、研究コミュニティによる最先端AI利用を支援するプロジェクト	オランダにおける計算・データ・研究情報基盤を統合提供する全国研究プラットフォーム
概要 (主体・予算・期間)	<ul style="list-style-type: none"> 欧州パートナーシッププログラムに最大4億9000万ユーロ、欧州連合以外のパートナーより最大5億ユーロの拠出を構想 	<ul style="list-style-type: none"> 2022年9月～2025年8月 (プロジェクト終了後も利用可能) 500万ユーロ (計算ノードや認証はEOSCと共有) 	<ul style="list-style-type: none"> オランダの教育・研究機関の非営利の協同体。 1億1694万ユーロ (2024年の実績ベースの収益)
システム概要	<ul style="list-style-type: none"> フェデレーション型メッシュアーキテクチャ。 	<ul style="list-style-type: none"> 複数の計算資源提供者を統合し、単一プラットフォームとして提供するPaaS。 	<ul style="list-style-type: none"> ストレージ基盤とデータ管理サービスは、SURFが内部で運用する共有インフラ上で提供。データセンターは全て国内に存在。
政策	<ul style="list-style-type: none"> 欧州研究領域 (ERA) の推進パイロット事業。 欧州共通データスペース (CES) の「研究・イノベーション」に属する。 	<ul style="list-style-type: none"> 研究インフラ(EOSC)上でのAIサービス提供が目的。 	<ul style="list-style-type: none"> オープンサイエンスNL実施計画2024-2025。 オープンサイエンス2030戦略。
AI活用	<ul style="list-style-type: none"> AI4EOSCほか 	<ul style="list-style-type: none"> プラットフォーム内でLLM/MLが利用可能 	<ul style="list-style-type: none"> 大規模AI基盤 オランダ語の国産AIモデル 等
学術知識グラフ活用	<ul style="list-style-type: none"> OpenAIRE PROVIDE RDGraph 		<ul style="list-style-type: none"> OpenAIREグラフを参照。
外部連携・留意点・課題	<ul style="list-style-type: none"> EOSC Interoperability Framework 	<ul style="list-style-type: none"> 必要に応じて、外部のデータセット、AIカタログ、ストレージ、他プラットフォームと連携し、EOSC内外の資源を取り込む。 	<ul style="list-style-type: none"> SURF EOSCノードとしてEOSCに参画。

調査結果概要 (2/2)

	NSF ACCESS (米国)	The American Science Cloud (米国)	DataON/ScienceON (韓国)
特徴	NSFが資金提供を行う先端計算・データリソースプログラム。	様々な分野にわたる科学研究、データ共有、および計算解析を促進・支援し、変革をもたらす人工知能モデルを実現するための、米国政府、学术界、および民間セクターのプログラムおよびインフラからなるシステム。	韓国科学技術情報研究院 (KISTI) が運営する国家研究データプラットフォーム。政府出資研究機関が生産する研究データを登録・共有・検索・利活用できる基盤で、国内外データ間の相互接続も進める。
概要 (主体・予算・期間)	<ul style="list-style-type: none"> 2022年5月～2027年4月 5200万米ドル (5年間) 2025年8月～2030年7月 (商用クラウド拡充) 2000万米ドル (5年間) 	<ul style="list-style-type: none"> 米国エネルギー省 (DOE) OBBB法第50404条で定義され、エネルギー省 (DOE) が管轄 2026年9月30日までに執行可能な\$150Mを予算配分 	<ul style="list-style-type: none"> KISTI (韓国科学技術情報研究院)
システム概要	<ul style="list-style-type: none"> ハイブリッド (全米の大学・研究機関が保有するオンプレHPCとNSF CloudBankを通じた商用クラウドを統合) 。 	<ul style="list-style-type: none"> インフラストラクチャーパートナーが提供する計算機資源をAmSCがオーケストレーションする構成 AmSCのサイエンスサービスとして、データ、AIスケール、モデルサービスとインテリジェントインターフェースを提供 	<ul style="list-style-type: none"> DataONはすべての実データを保有するわけではなく、研究機関が保有するレポジトリからメタデータをAPI経由で統合。
政策	<ul style="list-style-type: none"> CI・ブループリント (2019年) Genesis Mission (2025年) 	<ul style="list-style-type: none"> OBBB法第50404条 Transformational AI Models Genesis Mission (2025年) 	<ul style="list-style-type: none"> DataONは2018年に策定された国家戦略「研究データ共有・活用戦略」(科学技術情報通信部)
AI活用	<ul style="list-style-type: none"> NAIRR Pilot事業では、ACCESSの保有する既存資源・運用能力を中核として基盤が構築された 	<ul style="list-style-type: none"> 「システム概要」参照 	<ul style="list-style-type: none"> ScienceONにAI機能を組み込む 検索、翻訳・要約、比較分析
学術知識グラフ活用	<ul style="list-style-type: none"> Prototype Open Knowledge Network 		<ul style="list-style-type: none"> ScienceON
外部連携・留意点・課題			<ul style="list-style-type: none"> DataON/ScienceON (韓国)

海外プラットフォームの特徴と日本が取り組むべき具体策

海外プラットフォームから見える 次世代プラットフォームの特徴	日本の学術プラットフォーム整備において 取り組むべき具体策
<p>フェデレーション（連携）による資源の統合化</p> <p>既存のHPC、商用クラウド、研究データ、認証基盤を APIやメタデータ標準化を通じ、一つの巨大な仮想インフラとして統合・提供</p>	<p>計算基盤、データ基盤等の連携を実現し、分野や組織を問わず研究者がシームレスに活用できる環境を提供</p> <ul style="list-style-type: none"> • HPCや商用クラウド等の計算資源、データ基盤のオーケストレーション • 認証連携（シングルサインオン） • 計算資源を柔軟に利用を可能とする共通クレジット
<p>AI駆動型研究を支えるデータ・エコシステムの構築</p> <p>単なる研究データの保存に留まらず、AI基盤モデル、知識グラフによる高度検索、分析環境（Jupyter等）を一体的に提供</p>	<p>AIを専門としない研究者であっても容易にAI基盤モデルや計算資源を活用できる環境を提供</p> <ul style="list-style-type: none"> • FAIR原則推進によるAI-readyデータの整備 • 知識グラフの活用によるデータ検索・再利用機能の提供 • LLM等と連携したAI基盤モデル構築・利用を行える環境 • 研究者・研究機関への導入・活用支援
<p>非営利協同体・国主導による効率的な運営</p> <p>専門的な技術支援、トレーニングの提供、共同調達によるコスト効率化など、研究者が研究に専念できるサポート体制を組織化</p>	<p>プラットフォームの安定かつ持続的な発展・運営のための体制強化</p> <ul style="list-style-type: none"> • AI構築・活用を支援する人材の配置 • 産学官連携によるコミュニティ育成 • 財政基盤の強化

まとめ

データ創出、計算、データ・AIの3要素が密接に連携するエコシステムの構築が不可欠

- 実験機器等、計算資源と連携する、AI対応研究データ基盤およびネットワークを一体的に運用することで、データの安全な収集から高度な管理・活用までをシームレスに実現する情報基盤の整備が急務

研究分野およびHPCコミュニティからの多大な期待

- AIモデル・データ処理の高度化、研究・実務の自動化・支援、セキュリティ・インフラ環境、人材育成・知見の共有
- 実験装置、計算資源、データ基盤の高度な連携（資源オーケストレーション・認証連携）

海外動向への即応と国際競争力の維持

- 欧米の公的機関においても、計算・データ・認証基盤を統合した次世代プラットフォームの構築が同様の方向性で推進
- 海外動向と連携しつつ、国内にLLM構築等の専門知識・人材を蓄積することで、海外に後れを取らない自律的な研究基盤の確立が急務

AI for Science を支える研究データの管理・利活用と流通の在り方ワーキンググループ 今後の進め方及びとりまとめの方向性について（案）

資料6
AI for Scienceを支える研究データの管理・利活用と流通の在り方ワーキンググループ
(第5回)
令和8年5月22日

○これまでの議論

- 第1回 R7.12.24（水）
 - └ AI for Science を支える研究データの管理・利活用及び流通の在り方についての現状・背景等について文部科学省から説明
 - └ 学術研究プラットフォームの現状と今後について議論
- 第2回 R8.2.13（金）
 - └ SINETの現状と今後およびセキュリティについて議論
 - └ 「AI for Science の推進に向けた基本的な戦略方針について」の検討状況について文部科学省から説明および意見交換
- 第3回 R8.3.26（木）
 - └ 「AI for Science の推進に向けた基本的な戦略方針について」素案および具体的目標例について文部科学省から説明および意見交換
 - └ AI for Science における研究データの取り扱いに関する考え方について議論
 - └ 研究データ基盤の展開
- 第4回 R8.4.24（金）
 - └ AI for Science を支える研究データの管理・利活用及び流通の在り方ワーキンググループとりまとめに向けた整理について議論
 - └ 知識基盤の構築について議論
- 第5回 R8.5.22（金）
 - └ AI for Science を支える研究データの管理・利活用（主要分野の状況）について意見交換
 - └ AI for Science を支える研究データの管理・利活用及び流通の在り方ワーキンググループ（審議まとめ）素案について審議

○とりまとめに向けて

- 第6回 R8.6.16（火）
 - └ AI for Science を支える研究データの管理・利活用及び流通の在り方ワーキンググループ（審議まとめ）案について審議

AI for Science を支える研究データの管理・利活用と流通の在り方ワーキンググループ 審議まとめに向けて

○審議まとめ素案 構成

1. はじめに
2. 研究データの管理・利活用と流通を支える次世代情報基盤の必要性
 - (1) 背景 (AI for Science 推進とオープンサイエンス進展)
 - (2) 政策の方向性 (AI for Science の推進に向けた基本的な戦略方針、第7期科学技術・イノベーション基本計画等)
3. 学術研究プラットフォームの構築・実施状況
 - (1) 流通・ネットワーク基盤 (SINET 6、NII-SOCS) の現状
 - (2) 研究データ基盤 (NII RDC) の現状
4. 国内外の動向
5. ユーザーニーズ、社会からの要請
6. AI for Science の推進に向けた基本的な戦略方針への対応
7. 次世代情報基盤整備の方向性
 - (1) 付加価値の最大化を実現するネットワークの整備
 - (2) 研究データ管理・利活用の活性化、基盤高度化
 - (3) 計算資源との有機的な接続・連携、認証強化
8. まとめ

科学技術・学術審議会 情報委員会
AI for Science を支える研究データの管理・利活用と流通の在り方
ワーキンググループ運営規則

令和7年12月24日
科学技術・学術審議会情報委員会
AI for Science を支える研究データの管理・
利活用と流通の在り方ワーキンググループ

(趣旨)

第1条 科学技術・学術審議会情報委員会 AI for Science を支える研究データの管理・利活用と流通の在り方ワーキンググループ（以下「WG」という。）の議事の手続その他WGの運営に関し必要な事項は、科学技術・学術審議会令（平成12年政令第279号）、科学技術・学術審議会運営規則（平成13年2月16日科学技術・学術審議会決定）及び科学技術学術・学術審議会情報委員会運営規則（令和7年4月24日科学技術・学術審議会 情報委員会決定）に定めるもののほか、この規則の定めるところによる。

(議事)

第2条 WGは、当該WGに属する委員等の過半数が出席しなければ、会議を開くことができない。

- 2 WGの主査が必要と認めるときは、委員等は情報通信機器を利用して会議に出席することができる。
- 3 情報通信機器を使用した出席者は、第1項に規定する出席に含めるものとする。

(書面調査)

第3条 WGの主査は、やむを得ない理由により会議を開く余裕がない場合等においては、事案の概要を記載した書面等を委員等に送付し、その意見を徴し、又は賛否を問うことにより、書面調査を行うことができる。

- 2 前項の規定により書面調査を行った場合、WGの主査が次の会議において報告をしなければならない。

(会議の公開)

第4条 WGの会議及び会議資料は、次に掲げる場合を除き、公開とする。

- 一 WGの主査の職務を代理する者の指名その他人事に係る案件
- 二 行政処分に係る案件
- 三 前二号に掲げるもののほか、個別利害に直結する事項に係る案件、又は審議の円滑な実施に影響の生じるものとして、WGにおいて非公開とすることが適当であると認める案件

(議事録の公表)

第5条 WGの主査は、WGの会議の議事録を作成し、これを公表するものとする。

2 WGの会議が、前条各号に掲げる事項について調査審議を行った場合は、WGの主査が会議の決定を経て当該部分の議事録を非公表とすることができる。

(雑則)

第6条 この規則に定めるもののほか、WGの議事の手続その他WGの運営に関し必要な事項は、WGの主査がWGに諮って定める。

以上

海外公的機関におけるAI for Scienceを見据えた 研究インフラ（プラットフォーム）の動向

国立情報学研究所
2026年5月12日

概要

我が国のAI for Scienceを促進するための新たな情報基盤整備の参考とするため、海外の公的機関等におけるAI及び学術知識グラフを活用した学術基盤プラットフォーム構築事例に関する委託調査を以下の通り実施した。

■実施期間

- 2026年1月～3月

■調査項目

- 我が国での構築の参考とするため、学術基盤プラットフォームが、どのような政策を背景として、どのような機能・実装となっているか、AI及び学術知識グラフを活用しているかを把握する。

■調査対象

- 主要国の代表的なプラットフォーム10機関について、AI及び学術知識グラフの活用状況、公開情報の入手可能性、NII関係者のヒアリングも経て決定した。

■調査方法

- Web調査を主体とし、DataON/ScienceONについてはインタビューも実施して情報を補完した。

今回、上記調査のうちのEOSC(EU)、SURF(オランダ)、NSF ACCESS(米国)、DataON(韓国)についてNIIが編集・再構成した内容、及びNIIが追加で調査したThe AMERICAN Science Cloud(米国)を報告する。

調査結果概要 (1/2)

	EOSC(EU)	AI4EOSC(EU)	SURF(オランダ)
特徴	欧州の研究者がFAIR原則に基づき、分野や国境を越えて研究データを分散型で保存・共有・再利用できる、オープンな統合プラットフォーム	EOSC上でAI/ML/DL技術を活用するためのサービスを拡充し、研究コミュニティによる最先端AI利用を支援するプロジェクト	オランダにおける計算・データ・研究情報基盤を統合提供する全国研究プラットフォーム
概要(主体・予算・期間)	<ul style="list-style-type: none"> 欧州パートナーシッププログラムに最大4億9000万ユーロ、欧州連合以外のパートナーより最大5億ユーロの拠出を構想 	<ul style="list-style-type: none"> 2022年9月～2025年8月(プロジェクト終了後も利用可能) 500万ユーロ(計算ノードや認証はEOSCと共有) 	<ul style="list-style-type: none"> オランダの教育・研究機関の非営利の協同体。 1億1694万ユーロ(2024年の実績ベースの収益)
システム概要	<ul style="list-style-type: none"> フェデレーション型メッシュアーキテクチャ。 	<ul style="list-style-type: none"> 複数の計算資源提供者を統合し、単一プラットフォームとして提供するPaaS。 	<ul style="list-style-type: none"> ストレージ基盤とデータ管理サービスは、SURFが内部で運用する共有インフラ上で提供。データセンターは全て国内に存在。
政策	<ul style="list-style-type: none"> 欧州研究領域(ERA)の推進パイロット事業。 欧州共通データスペース(CES)の「研究・イノベーション」に属する。 	<ul style="list-style-type: none"> 研究インフラ(EOSC)上でのAIサービス提供が目的。 	<ul style="list-style-type: none"> オープンサイエンスNL実施計画2024-2025。 オープンサイエンス2030戦略。
AI活用	<ul style="list-style-type: none"> AI4EOSCほか 	<ul style="list-style-type: none"> プラットフォーム内でLLM/MLが利用可能 	<ul style="list-style-type: none"> 大規模AI基盤 オランダ語の国産AIモデル 等
学術知識グラフ活用	<ul style="list-style-type: none"> OpenAIRE PROVIDE RDGraph 		<ul style="list-style-type: none"> OpenAIREグラフを参照。
外部連携・留意点・課題	<ul style="list-style-type: none"> EOSC Interoperability Framework 	<ul style="list-style-type: none"> 必要に応じて、外部のデータセット、AIカタログ、ストレージ、他プラットフォームと連携し、EOSC内外の資源を取り込む。 	<ul style="list-style-type: none"> SURF EOSCノードとしてEOSCに参画。

調査結果概要 (2/2)

	NSF ACCESS(米国)	The AMERICAN Science Cloud(米国)	DataON/ScienceON(韓国)
特徴	NSFが資金提供を行う先端計算・データリソースプログラム。	様々な分野にわたる科学研究、データ共有、および計算解析を促進・支援し、変革をもたらす人工知能モデルを実現するための、米国政府、学术界、および民間セクターのプログラムおよびインフラからなるシステム。	韓国科学技術情報研究院(KISTI)が運営する国家研究データプラットフォーム。政府出資研究機関が生産する研究データを登録・共有・検索・活用できる基盤で、国内外データ間の相互接続を進める。
概要(主体・予算・期間)	<ul style="list-style-type: none"> 2022年5月～2027年4月 5200万米ドル(5年間) 2025年8月～2030年7月(商用クラウド拡充) 2000万米ドル(5年間) 	<ul style="list-style-type: none"> 米国エネルギー省(DOE) OBBA法第50404条で定義され、エネルギー省(DOE)が管轄 2026年9月30日までに執行可能な\$150Mを予算配分 	<ul style="list-style-type: none"> KISTI (韓国科学技術情報研究院)
システム概要	<ul style="list-style-type: none"> ハイブリッド(全米の大学・研究機関が保有するオンプレHPCとNSF CloudBankを通じた商用クラウドを統合)。 	<ul style="list-style-type: none"> インフラストラクチャーパートナーが提供する計算機資源をAmSCがオーケストレーションする構成 AmSCのサイエンスサービスとして、データ、AIスケール、モデルサービスとインテリジェントインターフェースを提供 	<ul style="list-style-type: none"> DataONはすべての実データを保有するわけではなく、研究機関が保有するレポジトリからメタデータをAPI経由で統合。
政策	<ul style="list-style-type: none"> CI・ブループリント(2019年) Genesis Mission(2025年) 	<ul style="list-style-type: none"> OBBA法第50404条 Transformational AI Models Genesis Mission(2025年) 	<ul style="list-style-type: none"> DataONは2018年に策定された国家戦略「研究データ共有・活用戦略」(科学技術情報通信部)
AI活用	<ul style="list-style-type: none"> NAIRR Pilot事業では、ACCESSの保有する既存資源・運用能力を中核として基盤が構築された 	<ul style="list-style-type: none"> 「システム概要」参照 	<ul style="list-style-type: none"> ScienceONにAI機能を組み込む 検索、翻訳・要約、比較分析
学術知識グラフ活用	<ul style="list-style-type: none"> Prototype Open Knowledge Network 		<ul style="list-style-type: none"> ScienceON
外部連携・留意点・課題			<ul style="list-style-type: none"> DataON/ScienceON(韓国)

EOSC

調査結果概要

FAIR原則に基づき研究データを分散型で保存・共有・再利用できるオープンな統合プラットフォーム

概要	<ul style="list-style-type: none"> 欧州の研究者がFAIR原則に基づき分野や国境を越えて研究データを分散型で保存・共有・再利用できる、オープンな統合プラットフォームとして開発中。既存の研究インフラを連携することでオープンサイエンスを推進し研究の生産性向上・透明性強化・再現性の確保を目指す。 	対象ユーザー	<ul style="list-style-type: none"> EUの研究者・技術者・教育関係者(4つのアクセスレベルを設定)
プロジェクト期間	<p>主要プロジェクト (Horizon Europeなどの枠組みで複数のプロジェクトが実施)</p> <ul style="list-style-type: none"> 2021年4月～2024年3月(EOSC Future) 2024年4月～2027年3月(EOSC Beyond) 	提供方法※5	<ul style="list-style-type: none"> アクセスグループにより、異なる認証方法でEOSC EUノードにログイン(学術機関:eduGAIN、EU職員:EU Login、EU加盟国の国民:eIDAS)
予算規模	<ul style="list-style-type: none"> 欧州委員会はEOSCのための欧州パートナーシッププログラムに最大4億9000万ユーロ、欧州連合以外のパートナーより最大5億ユーロの拠出を構想※1 上記うち主なプロジェクト予算 <ul style="list-style-type: none"> 約6,500万ユーロ(EOSC Future)※2 1,000万ユーロ(EOSC Beyond)※3 	アクセス制限※5	<ul style="list-style-type: none"> 有(所属機関またはIDプロバイダーから提供されるロール属性、および所属国に基づき、アクセスレベルを決定)
資金源	<ul style="list-style-type: none"> Horizon 2020/Horizon Europe(INFRAEOSC calls) 	利用料金※5	<ul style="list-style-type: none"> ログイン、閲覧は無料 各サービス(仮想マシン起動/実行中ワークフロー/ファイル同期・転送等)には、規模や使用期間に応じたクレジットコスト(仮想単位)を設定 サードパーティサービスは、プロバイダーにより別料金の場合あり
運営体制※4	<p>[所管・資金提供]</p> <p>EC(DG RTD+DG CNECT)</p> <p>EOSC-AISBL注</p> <p>[共同プログラムパートナー]</p> <p>[運営]</p> <p>EOSC Steering Board</p>	システム構成	<ul style="list-style-type: none"> 複数の組織に分散ノードを持つ、フェデレーション型クラウド
		研究データ保存を目的としたストレージまたは計算機の提供	<ul style="list-style-type: none"> 有り

注: AISBL(Association Internationale Sans But Lucratif)はEOSCを管理するため、2020年7月に設立された法的主体で、欧州委員会・加盟国と共同プログラム型パートナーシップを形成する。

*1 EC, European Open Science Cloud (EOSC) Partnership - MoU, https://research-and-innovation.ec.europa.eu/system/files/2022-01/c_2021_4113_f1_annex_en_v3_p1_1213802.pdf [2026/3/30閲覧]

*2 EC, Implementing the European Open Science Cloud, <https://digital-strategy.ec.europa.eu/en/funding/implementing-european-open-science-cloud> [2026/2/17閲覧]

*3 EOSC Beyond, About, <https://www.eosc-beyond.eu/about> [2026/2/17閲覧]

*4 RICH Europe, EOSC Governance structure, strategy and evolution, <https://rich-europe.eu/wp-content/uploads/2024/05/EOSC.pdfabout> [2026/2/17閲覧]

*5 EOSC EU Node, Access Policy, <https://open-science-cloud.ec.europa.eu/about/access-policy> [2026/2/17閲覧]

FAIR原則に基づく分散・フェデレーション型研究基盤

FAIR原則に基づき、研究データを分散型のまま保存・共有・再利用できるオープンな統合プラットフォーム

- **フェデレーション化**により、既存の研究基盤を置き換えることなく、相互運用性を確保、他ノードのリソース利用が可能に。
- サービスが独立・孤立していた非フェデレーション環境を、**認証(AAI: Authentication and Authorisation Infrastructure)**や**機能の統合により、ノード間でのシームレスなアクセス**を実現。

■ フェデレーション・アーキテクチャ

自律的な「EOSCノード(国・地域・テーマ別等)」のネットワーク化

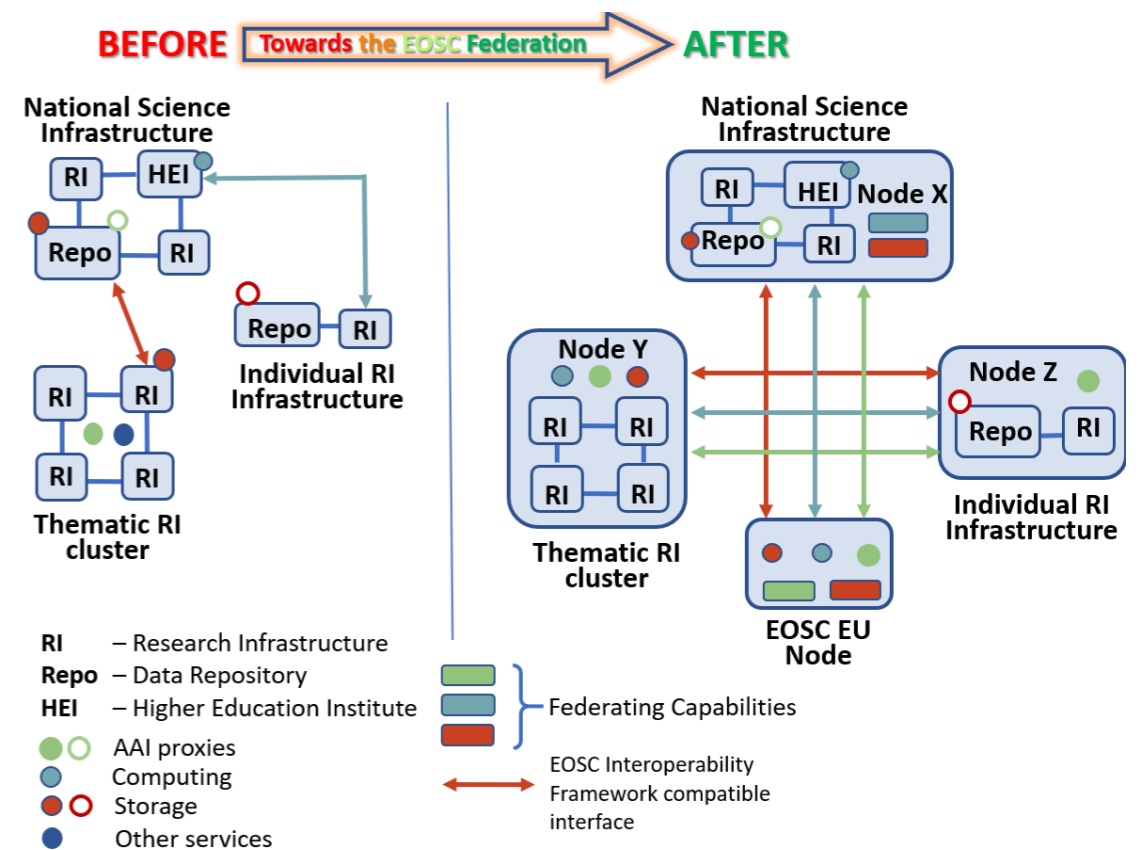
- フェデレーションサービス: エンドユーザーがデータやリソースを活用するための共通機能群。
- インターフェース標準化: APIやメタデータスキーマを用い、EOSC相互運用性フレームワーク(IF)に準拠することで、異なるノード間のサービスを接続。

■ EOSC EUノード: 最初の運用拠点/ハブ

EOSCフェデレーションにおける最初の欧州レベルの運用ノードであり、他ノードの基準点

- 研究者向けフロント: 計算資源(Compute)、ストレージ、ノートブック、ツール、カタログへの入り口を提供。
- 候補ノード向け: AAIやNode Registryなどのコアサービスを提供し、フェデレーション全体の相互運用性を支える。

研究サービスへのEOSCのインパクト



*1 EOSC Federation Handbook, <https://zenodo.org/records/18454649> [2026/3/24閲覧]

システム構成

1. フェデレーション型メッシュアーキテクチャ

自律的なノードのネットワークとして実装された「systems of systems(システムの連合体)」アーキテクチャを採用

- **相互運用メカニズム:** 各ノードが能力(Capabilities)を公開し、Node RegistryやAPIを通じて直接リソースを交換
- **コア機能の統合:** AAI(認証)、PID(識別子)、Resource Catalogue(メタデータ管理)を統合し、巨大な知識グラフ(OpenAIRE Graph連携)を構築。

2. 各国・地域の異種IaaS基盤を統合し、科学ミッションに対応する分散型研究インフラ

各国・地域の異なるIaaS基盤を統合し、特定の科学ミッションに対応する分散型インフラを実現

- EGI(European Grid Infrastructure) Foundationは、異機種混在のIaaSクラウドのフェデレーションとして動作。オープンスタンダードの採用により、ハイブリッド環境において、特定のクラウド基盤に依存せず、サービス移転が可能。
- **クラウドフェデレーション:** (i) フェデレーテッドAAI(研究者の本人認証とアクセス権限管理を行う仕組み)、(ii) フェデレーテッドアカウントिंग、(iii) 情報システム、(iv) フェデレーテッドモニタリング、(v) フェデレーテッドサービスレジストリ等のサービスで実現。

3. ownCloudを基盤とするクラウドストレージサービスとVirtual Machines基盤の計算機

研究者が安全かつ柔軟にデータ駆動型研究を遂行するための環境を提供

- **ストレージ:** ownCloud基盤のGDPR準拠クラウドストレージ。デバイス間の同期やリアルタイム共有が可能。
- **計算環境:** 拡張性の高いVirtual Machines(仮想マシン)と、Kubernetesベースのコンテナプラットフォームを提供。
- **HPC連携:** 欧州の高性能HPC基盤(PRACE等)と連携し、高度なAI解析やシミュレーションに対応。

4. クラウド/HPCノードの統合とパブリッククラウドによる効率化

既存の資産を活かしつつ、クラウドの柔軟性により効率を最大化

統合方式: 新たに物理装置を設置するのではなく、既存の大学・研究機関のデータセンター(クラウド/HPCノード)を論理的に統合。

効率性と持続可能性: 需要に応じた柔軟なリソース割り当てにより、待機時間を短縮。最新ハードウェアへの更新が容易なパブリッククラウドの利点を活かし、エネルギー効率も向上。

AIおよび学術知識グラフの利用

1. AIによる研究の高度化

■ AI4EOSC & AIoD

- EOSC上で高度なAI/ML/DLの開発・運用を可能にするプラットフォームを構築し、信頼性の高いAI環境を提供。
- 先進的な学習モデル：連合学習(Federated Learning)に焦点を当て、分散学習やサーバーレスコンピューティングをサポート。
- 信頼性の担保：MLOps(継続的開発・運用)の導入、ドリフト検出、豊富なプロビナンスメタデータにより、公平かつ堅牢なAIモデルを実現。
- AIoD(AI on Demand)連携：高性能画像解析サービス基盤の統合や、LLM/RAGを用いた高度な検索・推薦機能の実装。

■ EOSCエコシステムにおけるAI対応と自動化の実現(HORIZON-INFRA 2025年公募)

- AIによるメタデータの自動生成やデータ品質管理を行い、国内外のデータセットへのアクセスを最適化、FAIR化を促進：
- データ、計算資源、AIモデル、専門知識を統合し、エコシステム全体でのAI活用を支援する中核組織「EOSC AI/ML ケイパビリティセンター」を設立

2. 学術知識グラフ

■ OpenAIRE PROVIDE

リポジトリやデータアーカイブ、ジャーナル等の学術コンテンツをOpenAIREに接続するためのコンテンツゲートウェイサービス

- データソースを登録することで、研究成果はOpenAIRE GraphおよびEOSCエコシステムに統合され、EOSCノード管理者はダッシュボードで統合状況を監視できる。
- 構築されるナレッジグラフは、研究成果・助成金・組織情報を結び付け、検索、研究評価、資金提供者報告、FAIR/オープンサイエンスのモニタリングに活用される。

■ RDGraph (EOSC Research Discovery Graph)

EOSC内の研究発見のためのナレッジグラフ。

- EOSCナレッジグラフの拡張として構築され、EOSCの包含基準を満たす研究成果をOpenAIRE Graphから取得。EOSC Catalogueを拡張、研究者・分野・地域・資金・RAiD(研究プロジェクト識別子)・サービス等を統合しており、自然言語処理(NLP)とグラフベースAIで高度な検索を可能とする。

SURF

調査結果概要

SURFはオランダにおける計算・データ・研究情報基盤を統合提供する全国研究プラットフォーム

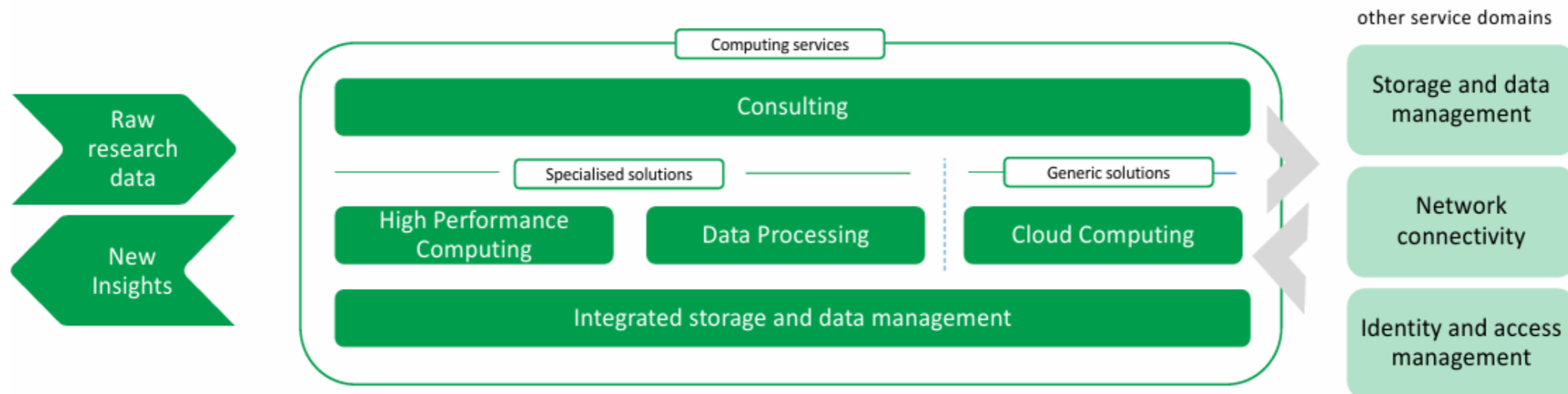
概要	<ul style="list-style-type: none"> SURFはオランダの教育・研究機関の非営利の協同体として、国家スーパーコンピュータSnelliusや研究者向けのResearch Cloud、永続的識別子RAiD等を通じ計算・データ・研究情報基盤を統合提供する全国研究プラットフォーム。
プロジェクト期間	<ul style="list-style-type: none"> 1987年に財団として設立、2015年に協同組合に、2020年にone SURFへ統合 Open Research Information Program(2025年7月～2030年12月)が主要な基盤構築プロジェクトとして進行中
予算規模	<ul style="list-style-type: none"> €116,941,000 (2024年の実績ベースの収益)
資金源	<ol style="list-style-type: none"> メンバーからの出資・会費・サービス利用料 政府・公共機関からの助成金・補助金 共同プロジェクト・外部資金
運営体制	<p>[運営]</p> <p style="text-align: center;">SURF</p>

対象ユーザー	<ul style="list-style-type: none"> 研究大学(WO)、応用科学大学(HBO)、中等職業教育機関(MBO)、大学医療センター(UMC's)、研究機関、その他(非営利系・非大学系研究機関(NOW, KNAW等))
提供方法	<ul style="list-style-type: none"> 会員向けサービスパッケージ 基本サービスパッケージ+セクター別サービスパッケージ 非会員向けサービスパッケージ(以下2つを要購入) 基本サービスパッケージ(インフラ)+(調達およびデジタルプラットフォーム) オプションサービス [詳細後述]
アクセス制限	<ul style="list-style-type: none"> 無し
利用料金	<ul style="list-style-type: none"> SURFのサービスは、会員向けと非会員向けの2つの異なるサービスパッケージと料金で提供されている。*1 会員向けのサービスは基本パッケージと、5つのセクターに応じて設定されるセクター別パッケージにより構成されている。 <p>*1 https://www.surf.nl/files/2025-10/sf_diensten26_en_web_v2.pdf</p>
システム構成	<ul style="list-style-type: none"> ハイブリッド型
研究データ保存を目的としたストレージまたは計算機の提供	<ul style="list-style-type: none"> 有り

プラットフォーム概要

SURFはオランダ国内の100を超える教育・研究機関が参加する共同体で、8つのサービスを提供 高性能な計算資源の提供に加え、専門的な助言やトレーニングも実施

- オランダの教育・研究機関の協同組合として、国家スーパーコンピュータSnelliusや研究者向けのResearch Cloud、永続的識別子RAiD等を通じ計算・データ・研究情報基盤を統合提供する全国研究プラットフォーム*1。オランダの100を超える教育・研究機関がSURFに参加し、限られたリソースや専門知識を共有し、重複投資を避け、イノベーションを効率的に進め、資源を有効活用している。
- SURFでは、信頼性の高い最新の設備やサービスを共同開発または共同調達。
- 提供するサービスは、教育・研究分野の共通ニーズに応える8つのサービス領域(①認証・認可(IAM)、②研究ネットワーク接続基盤、③研究基盤のセキュリティ、④計算基盤、⑤ストレージ基盤とデータ管理、⑥研究成果の公開、⑦柔軟な教育提供、⑧調達とサービス提供)。
- 国内の研究者に対して、コスト効率が高く、アクセスしやすい大規模な計算サービスを提供している。また、単なる計算資源の提供にとどまらず、専門的な助言やトレーニングも実施、これらのサービスは教育目的でも活用可能とされている。



*1 SURF, "About SURF", <https://www.surf.nl/en/about>, [2026/2/12閲覧]

*2 SURF, (2025), "SURF services portfolio", <https://www.surf.nl/files/2025-07/surf-dienst domein visies-slidedeck-engels.pdf>, [2026/2/18閲覧]

SURFは高性能な計算資源の提供に加え、専門的な助言やトレーニングも実施

[計算基盤]*1

■ 高性能計算(HPC)

- オランダのスーパーコンピュータSnelliusと欧州のプレ・エクサスケール級スーパーコンピュータLUMI(Large Unified Modern Infrastructure)利用により、研究者は大規模で複雑な計算やAIモデルの学習を実行可能。これらのシステムには、多様な研究用途に対応した標準構成とソフトウェアが整備。
- 計算タスクは順次処理され、必要なときに強力な計算資源へアクセス可能。

■ データ処理

- SURFでは、科学観測機器などから生成される大規模データ処理向けに、Grid(欧州のフェデレーション型インフラの一部)とSpider(ローカル環境で柔軟にカスタマイズ可能)という高スループット計算(HTC)基盤を提供。

■ クラウド型計算環境(SURF Research Cloud)

- 大規模な計算資源を必要とせず、使いやすく柔軟な環境を研究者が求める場合のニーズに対応。計算・ストレージ資源は必要に応じて自由に拡張・縮小可能。

■ コンサルティング

- 可視化、機械学習、HPC(高性能計算)、HTC(高スループット計算)、クラウドサービス等の専門分野で研究者を支援するコンサルティングサービスを提供。

[ストレージ基盤とデータ管理]

SURFが内部で運用する共有インフラ上で、クラウドストレージとテープストレージの組合せで提供。全てのデータセンターはオランダ国内にあり、SURF自ら管理。

■ データ転送 | SURF Filesender

- 最大1TBまでの小規模から大規模なファイルを暗号化された状態で送信可能。外部からのアクセスは技術的に不可能。*2

■ 同期・共有 | SURF Drive、Research Drive

- SURF Driveは、個人用クラウドストレージサービスで、最大1TBまでのデータを安全に保存・管理、他のユーザーと共有可能。欧州のプライバシー法に準拠、データはオランダ国内で安全に保管され、第三者提供はない。*3
- Research Driveは、研究チームによる共同利用や、学内外の研究者・学生とのコラボレーションを目的として設計されたストレージサービス。*4

■ ストレージ | Data Archive、Object Store

- Data Archiveは、大規模データセットの長期保存を目的、ペタバイト級まで対応可能。テープインフラを用いたコールドストレージ。
- Object Storeは、多様な種類のデータ(小規模ファイルから大容量データまで)に対応した、スケーラブルかつ柔軟性の高いストレージサービス。保存するデータに合わせて制限なく増量可能SURFの原則に準拠しながら、商用サービスに代わる信頼性の高い選択肢を提供。

■ 研究データ管理(RDM) | Yoda Hosting、RDM Storage Scaleout

- Yoda Hostingは、あらかじめ定義されたワークフローとウェブポータルを備えた、研究データ管理(RDM)ソリューション。Yodaは、iRODSコンソーシアム(SURFもこのメンバー)が開発・保守するオープンソースのデータ管理ソフトウェアiRODS上で動作するアプリケーション。*5
- RDM Storage Scaleoutは、iRODSベースの独自データ管理システムを運用する機関向けに、SURFのストレージサービスとの統合を可能にする拡張型ストレージソリューション。

*1 SURF, (2025), "SURF services portfolio", <https://www.surf.nl/files/2025-07/surf-dienst Domeinvisies-slidedeck-engels.pdf>, [2026/2/18閲覧]

*2 SURF, (2025), "SURF services portfolio", <https://www.surf.nl/files/2025-07/surf-dienst Domeinvisies-slidedeck-engels.pdf>, [2026/2/18閲覧]

*3 SURF (SURF User Knowledge Base), "SURF Drive", <https://servicedesk.surf.nl/wiki/spaces/WIKI/pages/166559750/SURFdrive>, [2026/2/20閲覧]

*4 SURF (SURF User Knowledge Base), "Research Drive", <https://servicedesk.surf.nl/wiki/spaces/WIKI/pages/117178843/Research+Drive>, [2026/2/20閲覧]

*5 SURF (SURF User Knowledge Base), "Yoda Hosting", <https://servicedesk.surf.nl/wiki/spaces/WIKI/pages/19824782/Yoda+Hosting>, [2026/2/20閲覧]

サービス提供

SURFは1987年に財団として設立され、2015年に協同組合に移行し、2020年より統合的に運営「会員向け」と「非会員向け」の2つの異なるサービスパッケージ・料金によりサービスを提供

- SURFのサービスは、会員向けと非会員向けの2つの異なるサービスパッケージと料金で提供されている。^{*1}
- 会員向けのサービスは基本パッケージと、5つのセクターに応じて設定されるセクター別パッケージにより構成されている。

会員向けサービスパッケージ

- SURFの会員は、基本サービスパッケージ(Basic Services Package)及びセクター別サービスパッケージ(Sector Services Package)を利用する。
- 基本サービスパッケージには、すべての会員が共通の固定料金で利用するサービスが含まれている。これは、SURFの共同基盤を構成する標準的サービス群であり、会員全体で費用を分担する仕組みである。
- セクター別サービスパッケージは、特定の分野(例:大学、応用科学大学、研究機関等)において当該セクター内で共同選定されたサービスを、単一料金で利用する制度で、各セクターのニーズに応じて構成される。
- さらに、会員はこれらのパッケージに加え、追加(オプション)サービスを選択することができ、当該サービスについては別途定められたサービス料金が適用される。追加の各サービスの料金は原則、会員・非会員に関わらず同一となっている。

非会員向けサービスパッケージ

- 非会員は、以下のパッケージを購入することによりサービス利用が可能になる。
 - ① 基本サービスパッケージ(インフラ)
 - ② 基本サービスパッケージ(調達およびデジタルプラットフォーム)
- これらのパッケージ購入により、当該パッケージに含まれる関連サービスへのアクセスが付与される。また、非会員についても、別途定められたサービス料金により追加(オプション)サービスを個別に購入することが可能。

^{*1} SURF, "SURF Services and Rates 2026", https://www.surf.nl/files/2025-10/sf_diensten26_en_web.v2.pdf, [2026/2/13閲覧]

NSF ACCESS

調査結果概要

HPCから商用クラウドまで、先端計算資源へのアクセスを統合する全米研究計算基盤

概要	<ul style="list-style-type: none"> NSFが資金提供を行う先端計算・データリソースプログラム。研究者・教育者が無償で国の先端計算システム・サービスを利用できるよう、国のサイバーインフラのアクセス性を向上させ、全米の大学にあるシステムや研究コミュニティとの統合を促進する。 	対象ユーザー	<ul style="list-style-type: none"> 米国の研究者・教育者、学生、国際協力者
プロジェクト期間	2022年5月～2027年4月*1 2025年8月～2030年7月(商用クラウド拡充)*2	提供方法	<ul style="list-style-type: none"> 複数大学・研究機関の計算資源(スーパーコンピュータ、クラスタ等)を統合し、ポータルを通じて利用者に提供*3 研究者は、研究ニーズに応じてプロジェクトタイプを選択し、ACCESS IDで割当を申請すると、提案書の審査後にクレジットが付与され、クレジットを使用してリソースにアクセス*4
予算規模	<ul style="list-style-type: none"> 5,200万米ドル(5年間)*1 2,000万米ドル(商用クラウド開発、5年間)*2 	アクセス制限	<ul style="list-style-type: none"> なし(ただし、NSFポリシーに準拠し、国家によるテロリスト等は排除*5)
資金源	<ul style="list-style-type: none"> NSF Office of Advanced Cyberinfrastructure(OAC) 	利用料金	<ul style="list-style-type: none"> 無償(ただし、リソース毎にクレジットの設定有り)
運営体制	<pre> graph LR NSF[NSF] --- Ops[運営] Ops --- CMU[カーネギーメロン大学] Ops --- CU[コロラド大学] Ops --- NYCU[ニューヨーク州立大学] Ops --- UIU[イリノイ大学] NSF --- Fund[所管・資金提供] </pre>	システム構成	<ul style="list-style-type: none"> ハイブリッド(全米の大学・研究機関が保有するオンプレHPCとNSF CloudBankを通じた商用クラウドを統合)
		研究データ保存を目的としたストレージまたは計算機の提供	<ul style="list-style-type: none"> 有り

*1 NSF, NSF ACCESS awardees will advance innovations in cyberinfrastructure accessibility, user support and integration services, <https://www.nsf.gov/news/nsf-access-awardees-will-advance-innovations> [2026/2/20閲覧]

*2 NSF, NSF expands access to advanced cloud computing for scientific research, <https://www.nsf.gov/news/nsf-expands-access-advanced-cloud-computing-scientific> [2026/2/20閲覧]

*3 ACCESS Operations, ACCESS Infrastructure Integration and Operations, <https://operations.access-ci.org/> [2026/2/20閲覧]

*4 ACCESS, For Researchers, <https://access-ci.org/get-started/for-researchers/> [2026/2/20閲覧]

*5 ACCESS Support, Update: ACCESS User Restrictions, <https://support.access-ci.org/announcements/update-access-user-restrictions> [2026/2/20閲覧]

プラットフォーム概要 | 研究データ保存を目的としたストレージまたは計算機の提供

オンプレミス中心の 計算リソース(HPC等)

- 大学・研究機関が分野特性を踏まえて整備したHPC群から構成
- 用途に応じた多様なアーキテクチャが併存

ストレージリソース

- 大学・研究機関が保有・運用するHPC環境に付随するストレージとして提供
- HPC運用に最適化されたストレージ構成

商用クラウドリソース Cloud Bank

- 商用クラウドを「ACCESSの割り当りリソース」として制度的に統合
- 割当は金額(ドル)ベースで実施し研究者自身が利用管理・責任を負う

Science Gateway

- GateWay運営者(PI)が割当を取得・管理
- 研究者自身が割当を申請する必要がなく、研究者・教育者・学生含むロングテールの研究者層も利用可能

オンプレ×商用クラウドのハイブリッド構成を構築

■ CPU計算

- Bridges-2(PSC):HPC+AI+データ解析の統合設計ディスク+テープの2層ストレージ(Ocean)と連携
- Expanse(SDSC):GPU/CPU混在型、データ集約型研究向け
- Stampede3(TACC):多様な研究分野に対応する汎用大型HPC など

■ GPU計算

- Delta(NCSA):ACCESSにおけるGPU計算の「基盤的・中核的」な計算資源
- DeltaAI(NCSA):DeltaをAI/ML特化・大規模モデル向けの拡張資源として補完・拡張
- Voyager(SDSC):科学・工学研究向けAI専用システム
- Neocortex(PSC):深層学習・グラフ解析を高速化する革新的AI基盤 など

■ 長期・大容量ストレージ基盤

- Granite(NCSA):テープアーカイブ300PB超の複製データ保管能力
- Ranch(TACC):HPC計算結果の長期保存向けディスク+テープ型アーカイブなど

■ 研究データ共有・配信基盤

- Open Storage Network(OSN):分散型データ共有・配信基盤研究機関間での大規模データ流通を支援 など

■ クラウド型計算環境

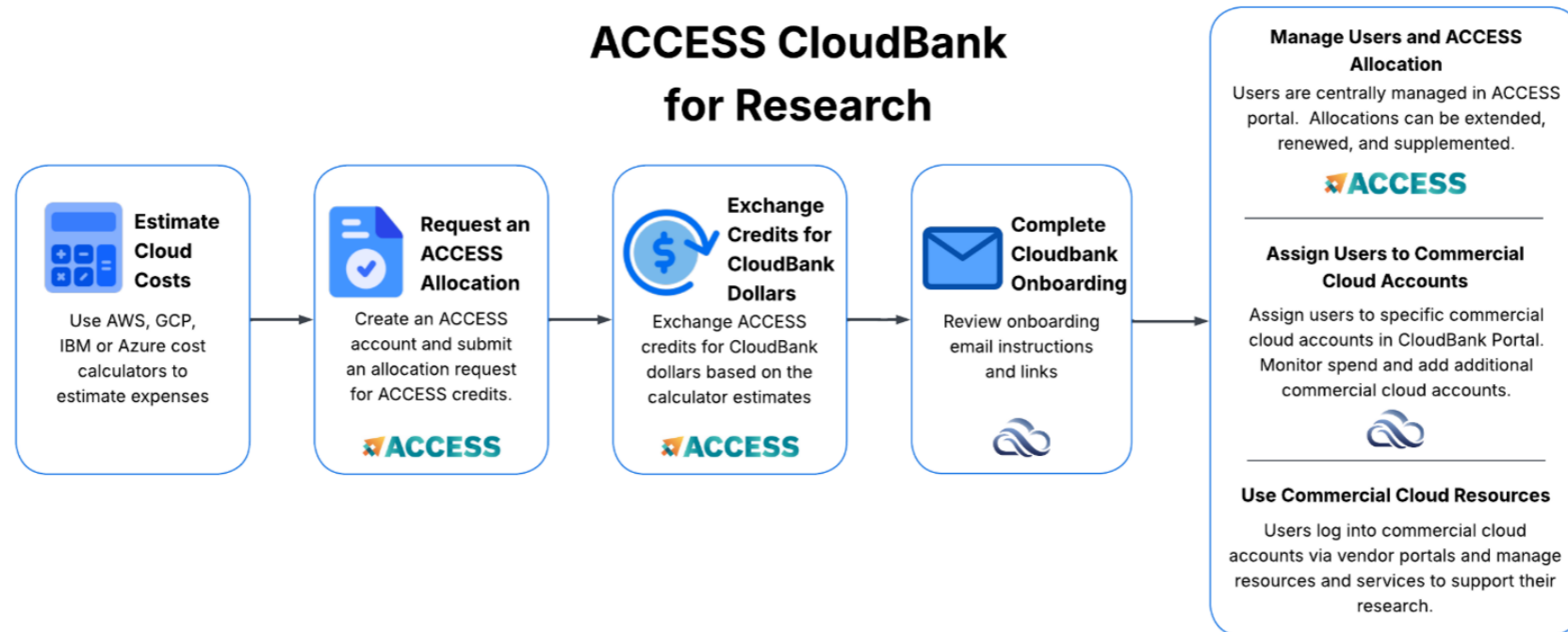
- CloudBank を通じて商用クラウド(Amazon Web Services、Google Cloud、IBM Cloud、Microsoft Azure等)を利用可能

Cloud Bank

商用クラウドを「NSFの研究基盤の一部」として制度化

- CloudBankは、ACCESSの割当制度を通じて、研究者が商用クラウドを金額ベースで利用できるようにするマルチクラウド仲介基盤。
- 商用クラウドをACCESSやNAIRR Pilotの割当プロセスに正式に組み込むことで、AWS、Google Cloud、IBM Cloud、Microsoft Azureを研究基盤の選択肢として位置付けることが可能となった。
- Cloud Bankでは割当が金額(ドル)ベースで管理されており、専用ポータルを通じて、利用料の追跡・支出通知・管理が行われる。
- オンプレHPCでは対応しにくい、AIの活用含む実験的な計算需要にも対応可能。

Cloud Bankにおける割当～利用までのプロセス



1)Cloud Bank, ACCESS Cloud Bank for Research, <https://www.cloudbank.org/training/access-cloudbank-research>, [2026/3/26閲覧]

国立スーパーコンピューティング応用センター(NSCA) Delta GPU

GPU計算・ストレージ基盤を全国で共有することでROI>1を達成

整備状況

■ システム構成*1

- A100 GPUノード×100
- A40 GPUノード×100
- Dense A100 GPUノード×5
- Dense MI100 GPUノード×1
- 各ノードに 1.6TB NVMe SSD(ローカルクラッチ)を搭載
- 200Gb/secのHPE Slingshot fabric により全 GPUノードおよび Delta ストレージに接続

■ 容量

- NCSA Delta GPU (Delta GPU) | Operations

■ 設置場所

- NCSA(University of Illinois Urbana-Champaign)

費用対効果

■ XSEDE/Jetstream/大学ローカルHPCのROI比較*2

- XSEDEは、NSFが11年間運用した共有CIプログラム(2022年8月終了)で、後継プログラムがACCESS。
- 3種類のサイバーインフラストラクチャリソース(XSEDE(共有CI)、Jetstream(クラウドシステム)、インディアナ大学Big Red II)の投資収益率を分析したところ、どれもROIは1超。
- 後続研究にて、連邦政府のXSEDEに対するROIは0.99から1.78に上昇(XSEDEの付加価値能力が時間とともに増加)し、全国共有モデルの費用対効果の高さが示された。*3

■ NSF大規模CI(XSEDE/Jetstream等)の実証データ*4

- NSFの資金提供を受けたXSEDEへの投資は総額約3億米ドル、創出された価値は47億~227億米ドル超と、費用対効果が高い(ROI>1を大きく超える)。

*1 ACCESS Operations, NCSA Delta GPU (Delta GPU), <https://operations.access-ci.org/node/593> [2026/2/20閲覧]

*2 Stewart, C. A., Hancock, D. Y., Wernert, J., Link, M. R., Wilkins-Diehr, N., Miller, T., ... & Snapp-Childs, W. (2018, December). Return on investment for three cyberinfrastructure facilities: a local campus supercomputer, the NSF-funded Jetstream cloud system, and XSEDE (the eXtreme Science and Engineering Discovery Environment). In 2018 IEEE/ACM 11th International Conference on Utility and Cloud Computing (UCC) (pp. 223-236). IEEE. [2026/2/20閲覧]

*3 Stewart, C. A., Costa, C. M., Wernert, J. A., Hancock, D. Y., McMullen, D. F., Blood, P., ... & Towns, J. (2022). Metrics of financial effectiveness: Return on investment in XSEDE, a national cyberinfrastructure coordination and support organization. In Practice and Experience in Advanced Research Computing 2022: Revolutionary: Computing, Connections, You (pp. 1-9).

*4 Spring Nature, "Evaluating Return on Investment for Cyberinfrastructure Using the International Integrated Reporting <IR> Framework", <https://link.springer.com/article/10.1007/s42979-024-02889-z> [2026/2/20閲覧]

The AMERICAN SCIENCE CLOUD

調査結果概要

Genesis Missionの主要な研究基盤としてDOE所有データによるAIモデルの構築と活用促進

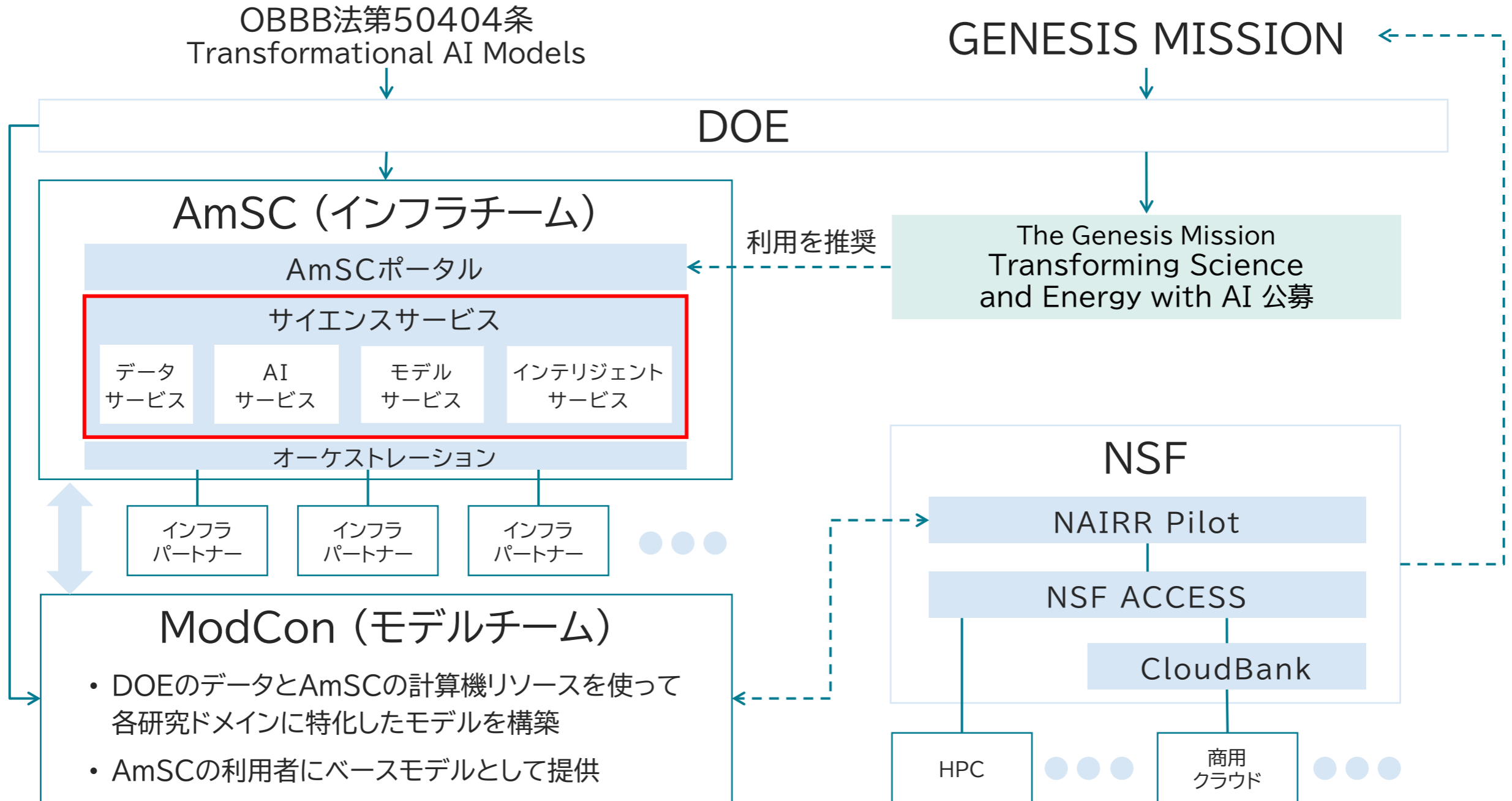
<p>概要</p>	<ul style="list-style-type: none"> 様々な分野にわたる科学研究、データ共有、および計算解析を促進・支援し、変革をもたらす人工知能モデルを実現するための、米国政府、学术界、および民間セクターのプログラムおよびインフラからなるシステム*1。 	<p>対象ユーザー</p>	<ul style="list-style-type: none"> DOEの科学者、施設利用者*3 The Genesis Mission: Transforming Science and Energy with AI 課題採択者*3 将来的には、より多くのDOEのデータ利用者*1
<p>プロジェクト期間</p>	<p>OBBB法第50404条で定義され、エネルギー省(DOE)が管轄*1</p>	<p>提供方法*3</p>	<ul style="list-style-type: none"> インフラストラクチャーパートナーが提供する計算資源やDOEのデータを、AmSCのポータル&ダッシュボードを介して提供。 インフラストラクチャーパートナーやサードパーティが提供するサービスやソフトウェアから、APIを通じて利用可能。 AmSC Identity Managementを介して利用可能。
<p>予算規模</p>	<ul style="list-style-type: none"> OBBB法第50404条を実行するために、2026年9月30日までに執行可能な\$150Mを予算配分(appropriation)*1。 AmSCを開発・導入を主導する統合チームの設立に向け、DOEが\$40Mの公募(LAB 25-3555)を実施(2025年12月に5件を採択)*2。 	<p>アクセス制限</p>	<ul style="list-style-type: none"> DOEの有するIDフェデレーション
<p>資金源*2</p>	<ul style="list-style-type: none"> DOE 民間セクターのプログラム(今後) 既存の計算基盤については、既存の予算 	<p>利用料金</p>	<ul style="list-style-type: none"> 無償
<p>運営体制</p>	<p>[所管・資金提供]</p> <ul style="list-style-type: none"> DOE [LAB 25-3555採択機関] <ul style="list-style-type: none"> オークリッジ国立研究所 フェルミ国立加速器研究所 ローレンス・バークレー国立研究所 トーマス・ジェファソン国立加速器施設 	<p>システム構成</p>	<ul style="list-style-type: none"> インフラストラクチャーパートナーが提供する計算機資源をAmSCがオーケストレーションする構成*1 AmSCのサイエンスサービスとして、データ、AIスケール、モデルサービスとインテリジェントインターフェースを提供
		<p>研究データ保存を目的としたストレージまたは計算機の提供</p>	<ul style="list-style-type: none"> インフラストラクチャーパートナーが提供*3 民間セクターのインフラ(将来)*2

*1 One Big Beautiful Bill (OBBB) Act (P.L. 119-21) <https://www.congress.gov/bill/119th-congress/house-bill/1/text> [2026/5/5閲覧]

*2 DOE funding announcement about The American Science Cloud (AmSC) <https://science.osti.gov/grants/Lab-Announcements/Lab-Announcements/2025/LAB-25-3555> [2026/5/5閲覧]

*3 AmSC Townhall: What is the American Science Cloud? https://amsc.energy.gov/wp-content/uploads/AmSC_for_RFA_Overview-040626.pdf [2026/5/5閲覧]

AmSCの構造とGENESIS MISSIONやモデル構築プログラムなどとの関係



AmSCのサイエンスサービスにおけるデータ & AI関連サービスの概要

データ関連サービスの概要

データ カタログ

- データセット検索のための集中型メタデータストア
- インフラパートナーや研究施設からのデータセット登録情報を統合したカタログ一覧

データ レイクハウス

- 高度にスケーラブルな SQL ライクなデータクエリをサポートする列指向データストレージ
- 生データから生成された、AI対応のキュレーション済みデータセット向けのホスティングソリューション

データ キュレーション

- 生データからAI対応データを処理・生成するためのパイプライン
- データレイクハウスへの取り込みの促進

データ 移動

- 登録済みのAmSCデータセットをインフラパートナーからコンピューティング施設へ転送
- 実験施設からコンピューティング環境へのストリーミングデータ転送

AI関連サービスの概要

大規模 サービス

- AmSCリソース全体にわたる厳選されたソフトウェアスタックを提供
- AIアプリのベンチマークを実施し、ソフトウェアスタックの機能とパフォーマンスの検証

モデルサービス

- MLFlow、ClearML、およびその他のフレームワークを活用した実験追跡とモデルカタログ
- AI向けDBaaS

推論サービス

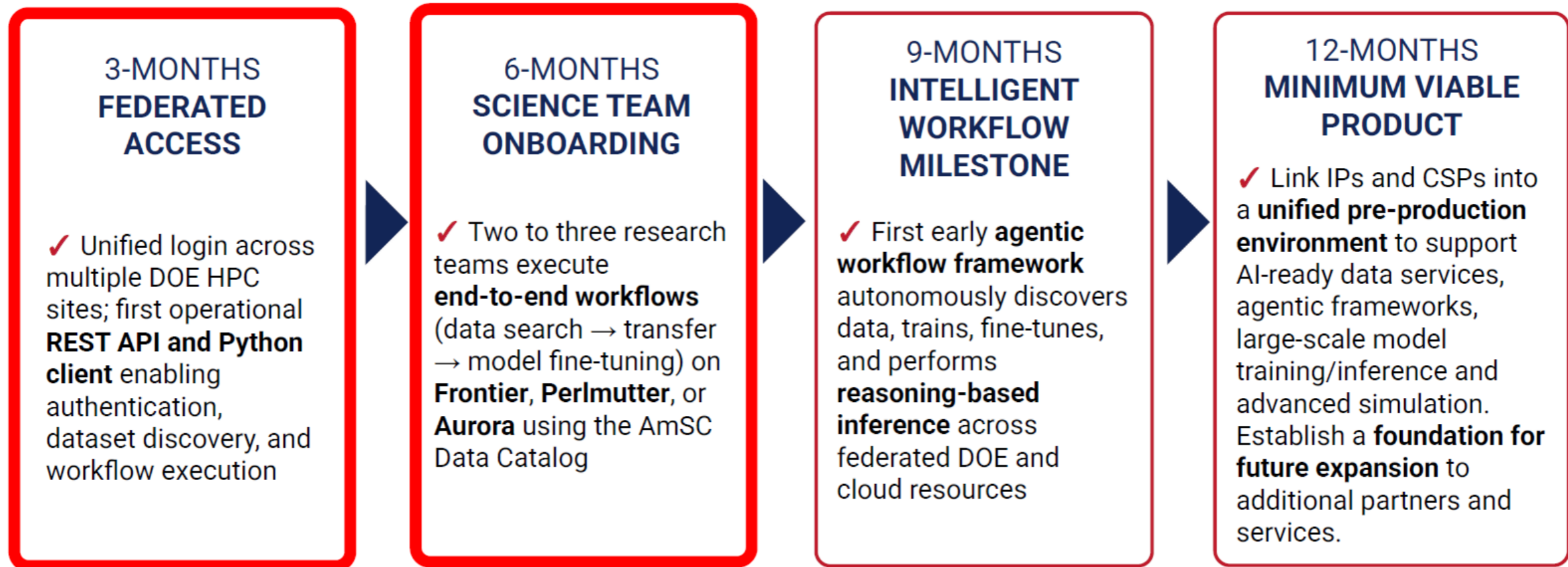
- DOEサイトおよび商用サービスにホストされたLLMおよび非LLMモデルの提供
- OpenAI準拠のAPIへのアクセス

インテリジェント インターフェース

- AmSCにおけるAgenticスタックの導入 (LLM APIへのアクセス、ChatUI、Agent Builder、MCPサーバーなどを含む)
- 施設および商用プロバイダーにおけるLLM推論サービスへのインターフェース

マイルストーン

- 第一の目標(3か月)として、DOEのHPC間でのID連携を実現。その後、APIを介してソフトウェアがデータセットの検索やワークフローの実行を実施できる環境を実装。
- 第二の目標(6か月)として、いくつかの研究チームが、AmSCデータカタログと指定するスパコンを使って、データ検索、データ転送、モデルのファインチューニングの実行を検証。
- 1年後には、将来の発展的な拡大に向けて、最低限の基盤連携のもとでAIを活用した一連の研究が実施できるAmSCの試行運用を開始。



KISTI DataON

調査結果概要

研究データの保存・分析機能を一体的に提供する国家研究データ基盤

概要	韓国科学技術情報研究院(KISTI)が運営する国家研究データプラットフォーム。政府出資研究機関が生産する研究データを登録・共有・検索・利活用できる基盤で、国内外データ間の相互接続も進める。
プロジェクト期間	2018年頃から構築開始、2020年1月よりサービス提供開始*1
予算規模	<ul style="list-style-type: none"> （立ち上げ当初の予算は不明） アップグレード予算は2021～2026年の72か月で251億ウォン（約21億円）*1 ハードウェア拡充費用を含む
資金源	<ul style="list-style-type: none"> 科学技術情報通信部からKISTIへの100%出捐*1
運営体制	<p>[所管・資金提供] [運営]</p> <pre> graph LR MSTI["MSTI (科学技術情報 通信部)"] --> KISTI["KISTI (韓国科学技術 情報研究院)"] KISTI --> DataON["DataON"] </pre>

対象ユーザー	<ul style="list-style-type: none"> 研究者(大学・政府機関・企業等) 政策立案者、一般利用者(市民科学者)
提供方法	<ul style="list-style-type: none"> Webポータルを通じた、検索・ダウンロード・登録機能 APIによる機械利用が可能 Virtual Research Environment(GUI型分析環境:β版) JupyterLab等)*2
アクセス制限	<ul style="list-style-type: none"> あり(会員登録に加えて、データ提供側がアクセスを制限できる)
利用料金	<ul style="list-style-type: none"> 無償(ただし計算資源を使用する場合は個別に費用発生)
システム構成	<ul style="list-style-type: none"> DataONはすべての実データを保有するわけではなく、研究機関が保有するレポジトリからメタデータをAPI経由で統合。
研究データ保存を目的としたストレージまたは計算機の提供	<ul style="list-style-type: none"> 研究データ保存機能あり

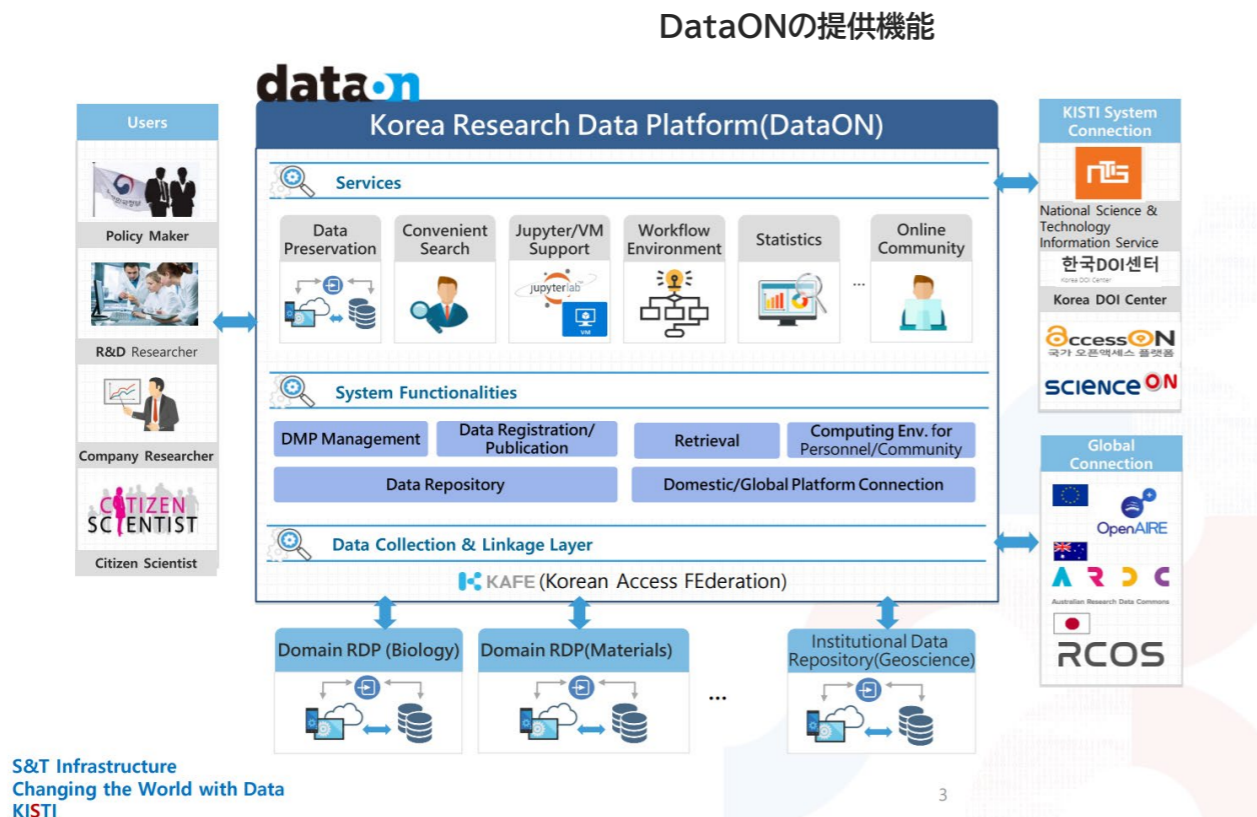
*1 KISTI, Establishing a System for Sharing and Disseminating Research Data, 2024, <https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO202500010709&dbt=TRKO> [2026/3/12閲覧]

*2 KISTI, Introduction to DataON, a national research data platform, <https://www.kisti.re.kr/promote/post/movie/4973>, [2026/3/2閲覧]

プラットフォーム概要

研究データプラットフォームとして保存から分析まで一体的に提供

- データ活用に係るサービスを一体的に提供
 - 利用者向け機能-基盤機能-データ収集・連携機能の3層構造で構成。
 - 各分野のデータ基盤やIDR、国家アーカイブと接続するハブとしてDataONは機能。



- 330万件*2
- 2,946,511データセット、1435ファイル*3
- メタデータ件数合計:約120万件(950GB相当)*4
- 図表データ:約560万件*4

KOREA RESEARCH DATA COMMONS, https://www.rd-alliance.org/wp-content/uploads/2024/05/KRDC_Introduction28GORC2928202207282C092CSa-kwang20Song29.pdf, [2026/3/2]閲覧

*2 KISTI, Introduction to DataON, a national research data platform, <https://www.kisti.re.kr/promote/post/movie/4973>, [2026/3/2閲覧]

*3 re3data.org, DataON Repository, <https://www.re3data.org/repository/r3d100013439>, [2026/3/2閲覧]

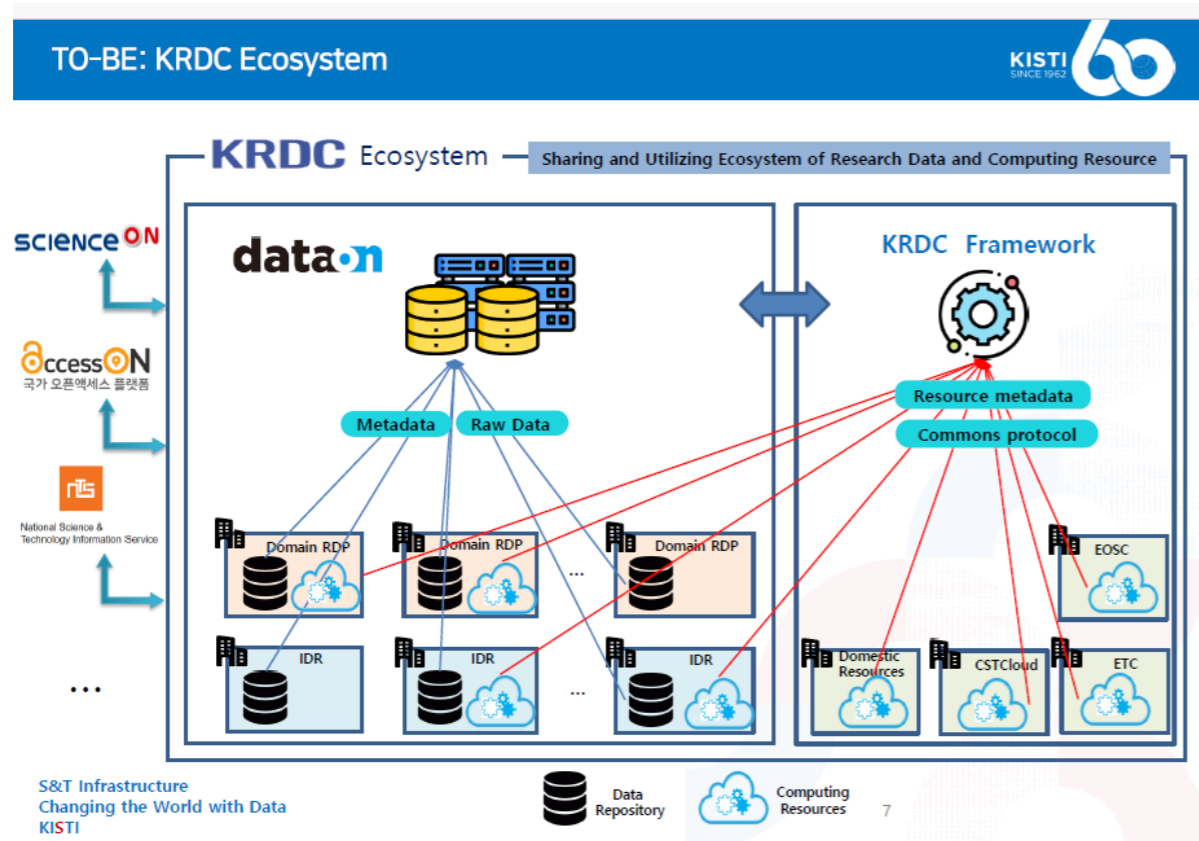
*4 KOREA RESEARCH DATA COMMONS, https://www.rd-alliance.org/wp-content/uploads/2024/05/KRDC_Introduction28GORC2928202207282C092CSa-kwang20Song29.pdf, [2026/3/2閲覧]

プラットフォーム構築の方針 | KRDC

2026年中をメドに、DataON・データレポジトリ・HPCリソースを結ぶKRDC構築を目指す

- KISTIが中心となって、国家研究データプラットフォームであるDataON、大学などに分散して存在する研究データレポジトリ、さらにはコンピューティングリソースを結び付け、リアルタイムに分析・活用できるKRDCエコシステムの構築を目指している。

KRDCエコシステム構想



KRDCのコア技術とフレームワーク

- ✓ KRDCは研究ソフトウェアと分析環境インフラを組合せて提供するサービスを目指しており、マルチノードKubernetesクラスタ基盤で動作。プロキシアプリ活用を想定している。
- ✓ KRDCフレームワークを活用してディープラーニング学習モデルの再現方法を提示し、ユーザーが望むデータとタスクを定義して学習モデルの再現と活用を支援する。
- ✓ コンピューティングリソースを管理するための標準メタデータスキーマを設計し、コンピューティングリソースの登録・検索・管理を簡素化する。

KOREA RESEARCH DATA COMMONS, https://www.rd-alliance.org/wp-content/uploads/2024/05/KRDC_Introduction28GORC2928202207282C092CSa-kwang20Song29.pdf, [2026/3/2閲覧]
 Design and Implementation of Workflow Federation Method for Multi-cluster Based Korea Research Data Commons,
<https://scienceon.kisti.re.kr/commons/util/originalView.do?cn=CFKO202333855028636&oCn=NPAP14389223&dbt=CFKO&journal=401134>, (2026/3/23閲覧)

まとめ

海外プラットフォームから見える 次世代プラットフォームの特徴	日本の学術プラットフォーム整備において 取り組むべき具体策
フェデレーション(連携)による資源の統合化 既存のHPC、商用クラウド、研究データ、認証基盤を APIやメタデータ標準化を通じ、一つの巨大な仮想インフラとして統合・提供	計算基盤、データ基盤等の連携を実現し、分野や組織を問わず研究者がシームレスに活用できる環境を提供 <ul style="list-style-type: none">• HPCや商用クラウド等の計算資源、データ基盤のオーケストレーション• 認証連携(シングルサインオン)• 計算資源を柔軟に利用を可能とする共通クレジット
AI駆動型研究を支えるデータ・エコシステムの構築 単なる研究データの保存に留まらず、AI基盤モデル、知識グラフによる高度検索、分析環境(Jupyter等)を一体的に提供	AIを専門としない研究者であっても容易にAI基盤モデルや計算資源を活用できる環境を提供 <ul style="list-style-type: none">• FAIR原則推進によるAI-readyデータの整備• 知識グラフの活用によるデータ検索・再利用機能の提供• LLM等と連携したAI基盤モデル構築・利用を行える環境• 研究者・研究機関への導入・活用支援
非営利協同体・国主導による効率的な運営 専門的な技術支援、トレーニングの提供、共同調達によるコスト効率化など、研究者が研究に専念できるサポート体制を組織化	プラットフォームの安定かつ持続的な発展・運営のための体制強化 <ul style="list-style-type: none">• AI構築・活用を支援する人材の配置• 産学官連携によるコミュニティ育成• 財政基盤の強化