

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

令和7年12月24日
(第1回WG資料)

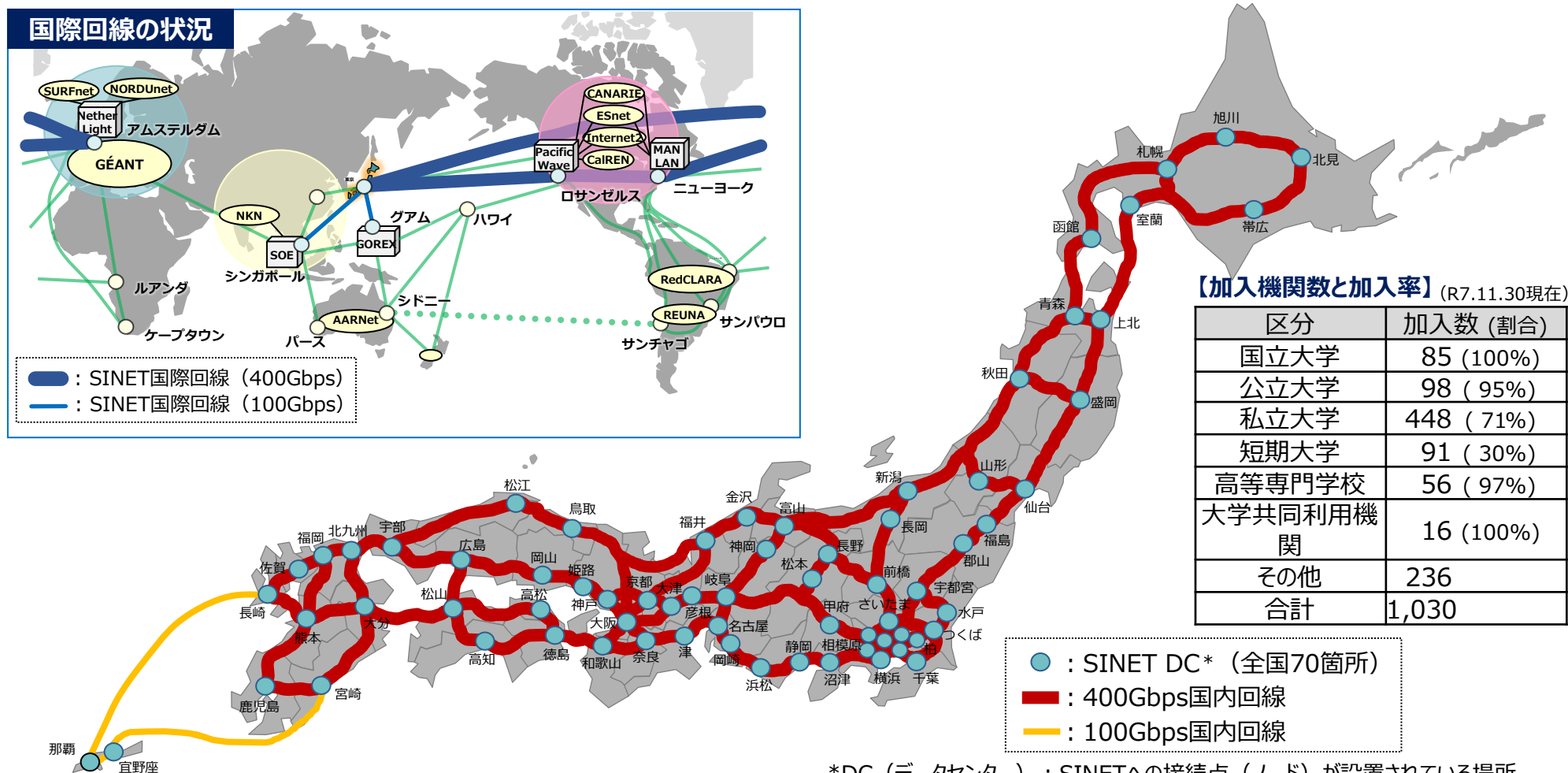
1. 現状・背景等

- ・学術情報ネットワーク（SINET）
- ・研究データ基盤（NII RDC）
- ・AI for Science を巡る動向等

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。

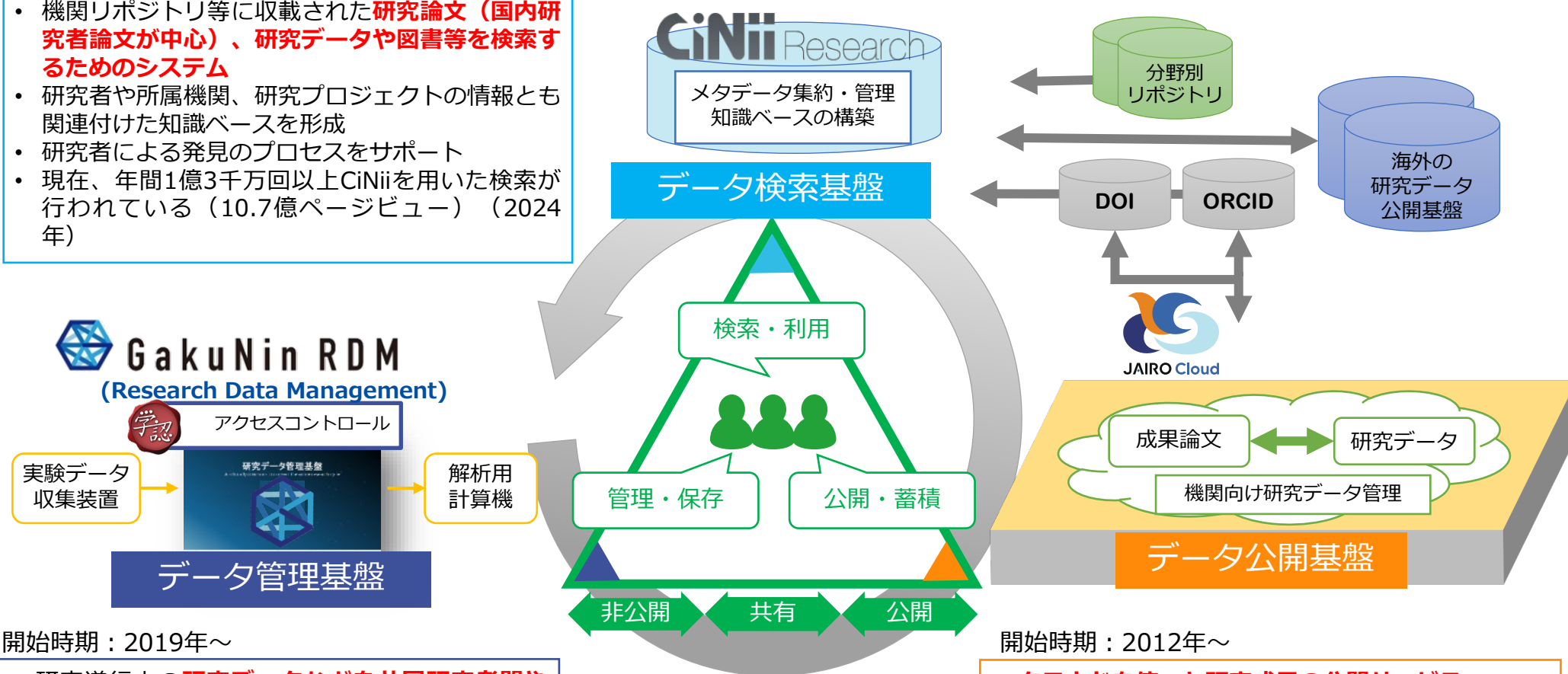
* NII RDC : NII Research Data Cloud



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載された**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）



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）

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

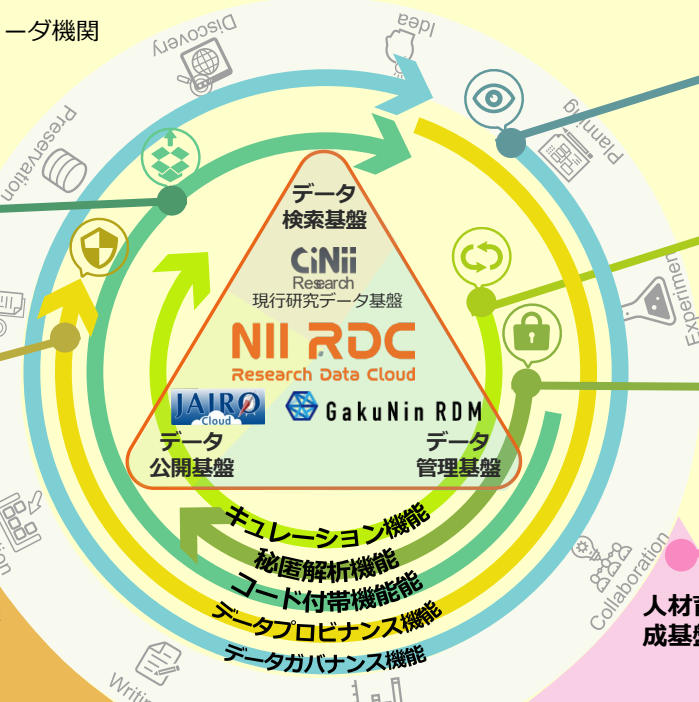
データ利用者がデータの来歴を確認できるとともに、データの提供者が自身のデータの利用状況を確認できる機能

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推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

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

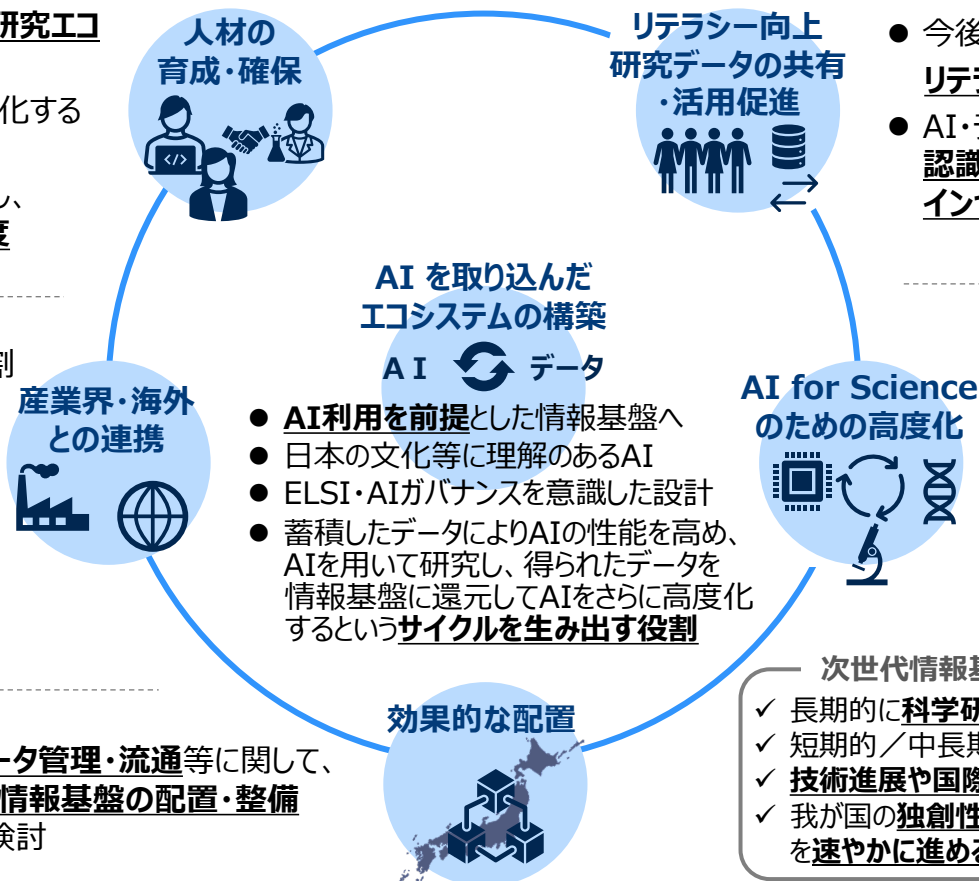
- 研究データ等の保存・管理、流通、活用を支える**情報基盤（※）は、AI時代の新たな科学技術・イノベーションを切り開くインフラ**となることが期待
※研究データ基盤、流通基盤、計算基盤をシームレスに接続した学術研究を支える基盤
- **AIを活用してあらゆる垣根を超えた新たな知の創造を支援し、AIが出力する情報の信頼性を担保する新たな情報基盤の構築が必要**
- 蓄積された多くの良質な研究データを学習データとして提供することで**AIモデルの更なる高度化、AI for Scienceの拡大、分野融合やすそ野の広い研究の促進、ひいては社会課題の解決や我が国全体の研究力・産業競争力の向上**に大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

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

- **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

- **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



- 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

- 「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**

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**

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

(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（ハイパフォーマンズ・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、AI 処理能力・アプリケーション実効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、AI 時代に適した研究データ基盤 NII RDC や流通基盤 SINET の高度化を推進する。

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、技術職員や URA等の人材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。
- ・上記共用拠点は、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、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の整備・充実等を進める。

AI for Science による科学研究の革新

- **日本固有の強み**を活かし、**ライフサイエンス**や**マテリアルサイエンス**をはじめとした分野横断的・組織横断的な取組を進めるとともに、**情報基盤**の強化や先端研究設備・機器の戦略的な整備・**共用・高度化、大規模集積**等を通じて「AI for Science」の先導的実装に取り組み、**科学研究システムを革新**する。

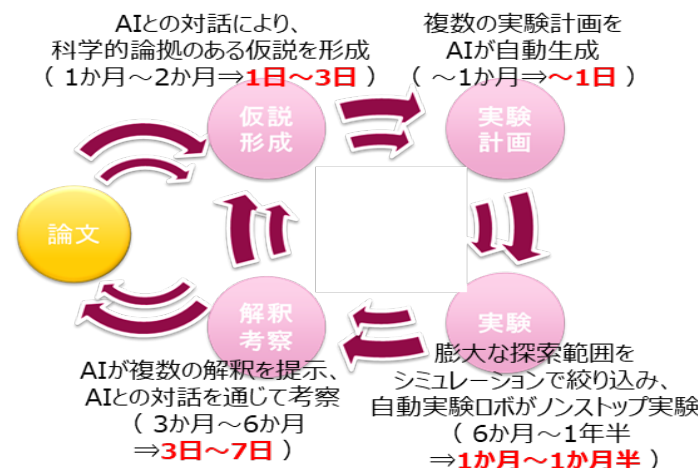
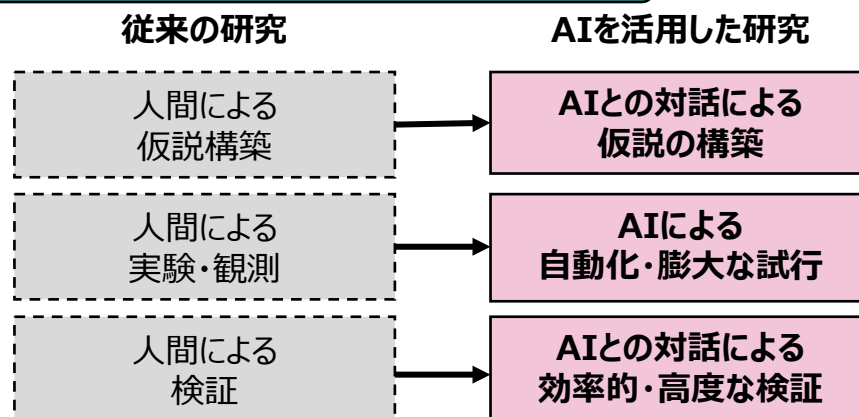
■（政策として）AI for Science による科学研究の革新とは・・・

- **AI技術を科学研究のあらゆる段階に適用し様々な分野で活用する取組とともに、AI研究、環境構築、人材育成、社会実装などを政策的に検討し、推進すること。**

- ・ AIが科学研究を高度化・高効率化すること
- ・ AIが科学研究を自律的に駆動すること
- ・ AIそのものの研究開発（Science for AI）
- ・ AI for Scienceを実現するための環境構築
- ・ 科学研究から社会実装への取組

多様な分野におけるAIの活用	活用例
科学研究で創出されるデータの改良や情報の抽出	医学領域における超音波画像診断支援／宇宙観測データのノイズ除去／古文書に記述されている内容の自動解析
シミュレーションの高度化・高速化	タンパク質の立体構造予測／気象予測／材料分野における望ましい特性を持つ材料や反応の発見／仏像の顔の類似度や制作年代・地域の推定
実験や研究室の自律化	自律的な物質探索ロボットシステム／抗体遺伝子クローニング(同じ遺伝子型となる細胞集団を作製すること)の自動化システム
新しい研究テーマ等の提案	研究データや論文情報の解析による科学的仮説の生成

AIによる研究の加速のイメージ



1. はじめに

- **AIは、研究力の生産性・効率性の向上のみならず、科学研究の在り方そのものを変革。**
経済成長、安全保障、地方創生、人手不足の解消、知の継承、災害への備えなどの社会課題の対応に不可欠。
- 海外では、**データ駆動型研究の推進や、研究設備の自動化・リモート化・自律化による大規模ハイスループット研究拠点の構築**により、**研究の高度化・高速化が急速に進展**。
- 日本では、イノベーションを促進しつつリスクに対応するため、本年5月に**AI法**が成立、9月に**AI戦略本部**を設置、本年冬を目前に**AI基本計画**が策定される予定。
『世界で最もAIを開発・活用しやすい国』を目指す方針。
- 研究活動におけるAI利活用（**AI for Science**）の急速な進展により、あらゆる分野で**研究生産性が飛躍的に向上**しようとしている。我が国もこの潮流に乗り遅れてはならず、**技術的優位性・不可欠性**を確保しなければならない。
- 日本の強みを活かした**「AI for Science」の先導的実装**に取り組むことが喫緊の課題。
- AIは、基礎研究段階を含めあらゆる分野の科学研究の姿を根本から変えつつある。AI for Science を我が国の**「科学の再興」の駆動力**とし、日本の科学力の**反転攻勢のチャンス**とするためには、**ここ数年が勝負、スピード感を持って取り組むことが必要**。

2. 日本の強み

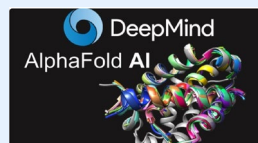
- 我が国は、日本全国をつなぐ流通基盤（SINET）や研究データ基盤（NII RDC）、世界有数の計算基盤（理研のもつAI for Science開発用スパコン、富岳、富岳NEXT（NVIDIA・富士通とともに開発中））など、**世界にも稀な世界最高水準の情報基盤を有している。**
- また、ライフサイエンスやマテリアル、防災、地球環境などの分野においてこれまで蓄積してきた**質の高い実験・観測データは、AI for Science推進のための大きな資産。加えて、数理科学・数理工学を始めとした基礎科学力の蓄積を保有。**
- さらに、日本は、世界有数の経済規模を持ち、社会的基盤が整っていることや技術水準の高さなどに加え、**AIやロボットに対する需要や社会的受容性が比較的高く、制度的にもAI導入に適した環境が整っている。**
- 長年培われてきた中小企業などにおける**先端機器等の製造能力**を保有。
- 我が国では世界に先駆けて**少子高齢化・人口減少が進展**。AI for Scienceの駆動力は、特定の分野で人間を超える処理能力をもつAIによるマンパワーの代替・拡張でもあり、人材不足等の課題を抱える日本において、**AI for Scienceの推進で「科学の再興」を目指す**ことが求められていると言えるのではないか。



3. 「AI for Science」の推進により目指す将来像

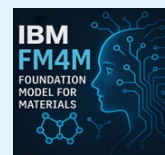
①「科学基盤モデル」の国産開発によるAI駆動型研究開発の強化 ②研究システムの自動・自律・遠隔化による研究データ創出・活用の高効率化

- ✓ バイオ分野の基盤モデルの開発により、複雑な生命現象の解明や、高精度な生体分子の構造予測、代謝・合成プロセスの予測等の効率化・最適化が可能になり、**バイオものづくりや医療・創薬研究のスピードを向上**
複数のモデルの組合せ等により、**仮想細胞モデルやデジタルツインを活用した、個別化医療を実現**



✓ 研究設備・機器の自動・自律・遠隔化のためのAI

- ✓ 膨大なマテリアル・データで学習した材料分野基盤モデルにより、これまでの限界を超えるような特性を持つ**革新的マテリアルの迅速な探索・開発が可能に**



✓ AI高度化に必要な良質かつ大量のデータ提供

✓ AIによる膨大なデータの管理効率化

✓ AI基盤モデルの構築・高度化に必要な計算資源・データの提供

③「AI for Science」を支える次世代情報基盤の構築

- ✓ より高度なAI基盤モデルの開発のためには、**膨大な計算資源や良質な研究データ**が不可欠。我が国には、研究データの管理・利活用の中核的なプラットフォームの研究データ基盤（**NII RDC**）や、日本全国の大学・研究機関等を超高速・低遅延でつなぎ、流通させる**SINET**、世界最高水準のスパコン「**富岳**」が存在。
- ✓ AI for Science 専用スパコンの運用や、「**富岳NEXT**」の開発・運用を通じて**AI処理能力・アプリケーション実効性能が飛躍**するとともに、国産技術が国際市場に訴求。
- ✓ **SINET**の高度化を通じて、**爆発的に増大し続けるデータ流通を安全かつ高速に支える**とともに、AIを活用した**NII RDC**の高度化を通じて、**研究データ管理等の研究者の負担となる業務を代替し、研究者の創造的活動の時間の確保**に貢献。



世界最高水準のAI・シミュレーション性能を目指す

2021年～

新たなフラッグシップシステム
（通称：富岳NEXT）

2030年頃までに運転開始



4. 基本的な戦略の考え方

- 日本の取るべき基本戦略としては、日本の資産・リソースを十分に活かして、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。
- AI for Scienceを推進するためには、それを支える**研究インフラ（情報基盤やデータ創出基盤等）の構築・整備が不可欠**。AIの利活用を前提に、分野特性や機関特性を踏まえた**俯瞰的視点で、横串を通す中長期的視点から、研究インフラと研究体制を一体的・戦略的に整備・構築**していく。
- その上で、日本の勝ち筋となり得る先導的分野等において、データ基盤の充実と分野特化型科学基盤モデルの開発等の先駆的取組の早期実装を通じ、世界のトップ水準に引き上げるとともに、次の種や芽を生み出す**萌芽的・探索的研究**を推進する。
- 加えて、AI時代における社会インフラとなるAIについては、他国に依存することなく、自国で研究開発する能力を保持することは安全保障上も極めて重要であり、**信頼できる（される）AIに関する取組や、AIそのものの研究開発（Science for AI）を持続的に強化**する。
- 併せて、AI関連人材の育成・確保を全てのレイヤーで推進する。高度な人材は高度な研究活動を通じて育成されるため、国内外の優秀な人材を引き付ける魅力的な研究環境を構築し、戦略的な国際連携により、研究レベルと人材レベルを世界トップレベルに追随していく。
- また、先駆的取組等を通じて、あらゆる分野においてAI for Scienceを波及・浸透させていくことで、**2030年代、全国どこでも誰でも、AIを使った研究活動が可能となる社会を実現**する。加えて、**科学とビジネスが近接化**しており、科学研究から**産業界への橋渡し**を通じて、産業界にもAI for Scienceを浸透させ、**科学とビジネスの好循環**を作り、**人口減少下における日本の労働力不足等の課題解決にも貢献**する。
- こうした取組を戦略的に推進するため、**国としての推進体制を構築し、研究インフラ・研究システムを抜本的に改革**する。

5. 今後の方向性と課題等

- **日本全体のAI for Science をスピード感を持って戦略的に推進**することが必要。国際的な潮流の中で、日本の強みを活かしてプレゼンスを示し、研究力を反転して行くためには、AIの利活用を前提に研究基盤・研究システムを転換し、研究活動におけるAI利活用により研究の効率性・生産性を向上して研究者の創造性を最大化していくとともに、AIそのものの研究等を推進することが必要。

① **AI研究（Science for AI）とAI利活用研究（AI for Science）における先駆的・先導的取組の推進**

- 先導的・先駆的取組の加速、萌芽的・探索的研究の推進
- 「AI駆動型研究開発」のきっかけ作り、波及・振興

② **AI駆動型研究を支えるデータの創出・活用基盤の整備**

- 研究システムの自動・自律・遠隔化による研究データの創出・集約・活用の高効率化

③ **AI for Science を支える次世代情報基盤の構築**

- 計算資源の拡張・共用、戦略的な研究資源の分配、リソースのマッチング
- 今後の研究データの管理・利活用・流通の在り方（次期SINETの検討など）

本WGでの
主ターゲット

④ **AI関連人材の育成・確保、産学・国際連携の強化**

- 人材育成・確保（トップサイエンティスト・データサイエンティストの育成確保、リテラシー向上、リ・スキリングなど）
- 産学連携・協働の推進（科学研究から産業への橋渡し、スタートアップ支援など）
- 国際連携・国際協調

⑤ **大胆な投資資金の確保・環境整備**

⑥ **推進体制の構築**

- 柔軟かつ機動的に研究開発を支援するための推進体制の構築
- 中核的な拠点のネットワーク化、コミュニティ（学会含む）の強化
- オープン＆クローズ戦略などデータ戦略や信頼性確保等に関する検討

AI for Scienceを支えるための情報基盤

- 研究データを実験等により創出した後、適切に保存・管理を行い、蓄えられたデータを流通させ、解析等により活用するといった一連の研究プロセスは、それぞれの役割を担う**研究データ基盤、流通基盤、計算基盤の3つの基盤**（総称して「**情報基盤**」と呼ぶ）が支えている。
- AI for Scienceの推進においては、**3基盤それぞれの高度化、及びこれらの基盤間の連携の強化等による次世代システムへのアップグレード**が必要。

研究データ基盤

- ✓ 様々な研究分野や機関の研究者が自身の**研究データを保存・管理、公開する中核的なプラットフォーム**（※）。
- ✓ 全国の図書や論文、根拠データなど研究に必要な**学術情報を一元的に検索し**、研究活動での利活用を実現。
- ✓ **オープンサイエンスやAI for Scienceの潮流を踏まえ、AI機能等を搭載**することにより、研究者にとってより**使いやすく、かつ、研究データが自然に集まってくるようなシステム**へと高度化を検討。

※「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決定）において、研究データ基盤（NII RDC）を中核的なプラットフォームとして位置付け



流通基盤

- ✓ 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**する通信ネットワーク（SINET）を整備。
- ✓ 実験機器や計算基盤から創出される**大容量データ**をはじめ、日々の研究活動から生じる研究データは、SINETを通じることで研究者・研究機関間の**高速・安全かつ遅延のない流通**を実現。
- ✓ データ駆動型の研究様式への進展に伴う**研究データの質・量の増大**を踏まえ、**より大容量の通信**に対応した次期ネットワークを検討。

計算基盤

- ✓ 世界最高水準の性能を有する**スーパーコンピュータ「富岳」**を中核とした、多様な利用者のニーズに応える**革新的な計算環境（HPCI）**を構築し、全国のユーザーへ計算資源を提供。
- ✓ **世界最高水準のAI・シミュレーション性能を有する新たなフラッグシップシステム**を2030年頃までに整備・運転開始予定。
- ✓ 生成AIをはじめとした技術革新に対応した次世代計算基盤を構築し、**成果創出のためのアプリケーション開発や環境整備**を実施予定。



活用

流通

【SINET加入機関数と加入率】(R7.11.30現在)

区分	加入数 (割合)
国立大学	85 (100%)
公立大学	98 (95%)
私立大学	448 (71%)
短期大学	91 (30%)
高等専門学校	56 (97%)
大学共同利用機関	16 (100%)
その他	236
合計	1,030

SINET6



2. 主な論点（案）

【流通基盤 SINET】

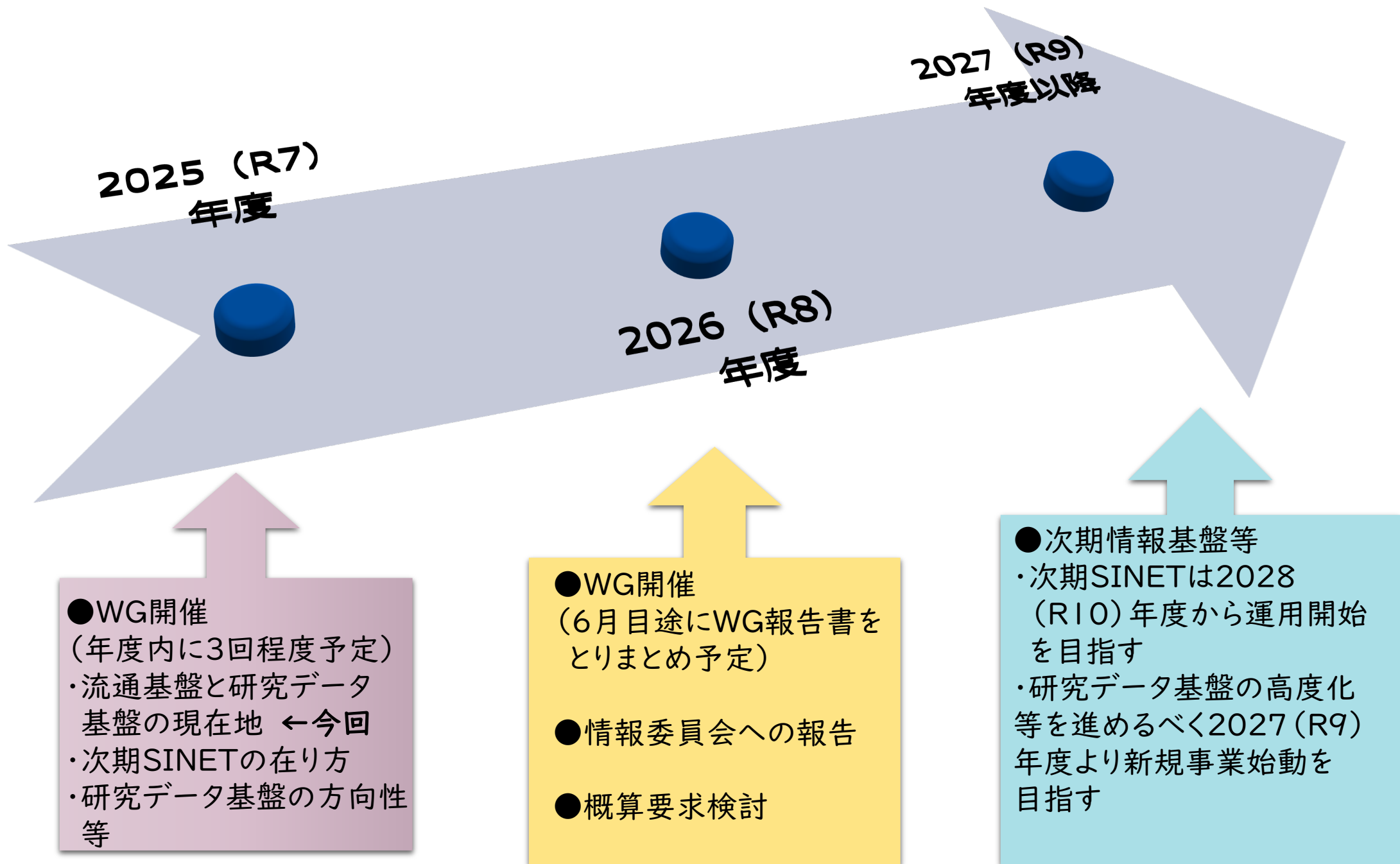
- 現行SINET（SINET6）の活用状況
- 次期SINETに向けたニーズの把握
- 最新の技術動向
- 海外の情報通信ネットワークの動向
- サイバーセキュリティへの対応

【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

3. 今後のスケジュール（案）

次期情報基盤の運用等に向けて

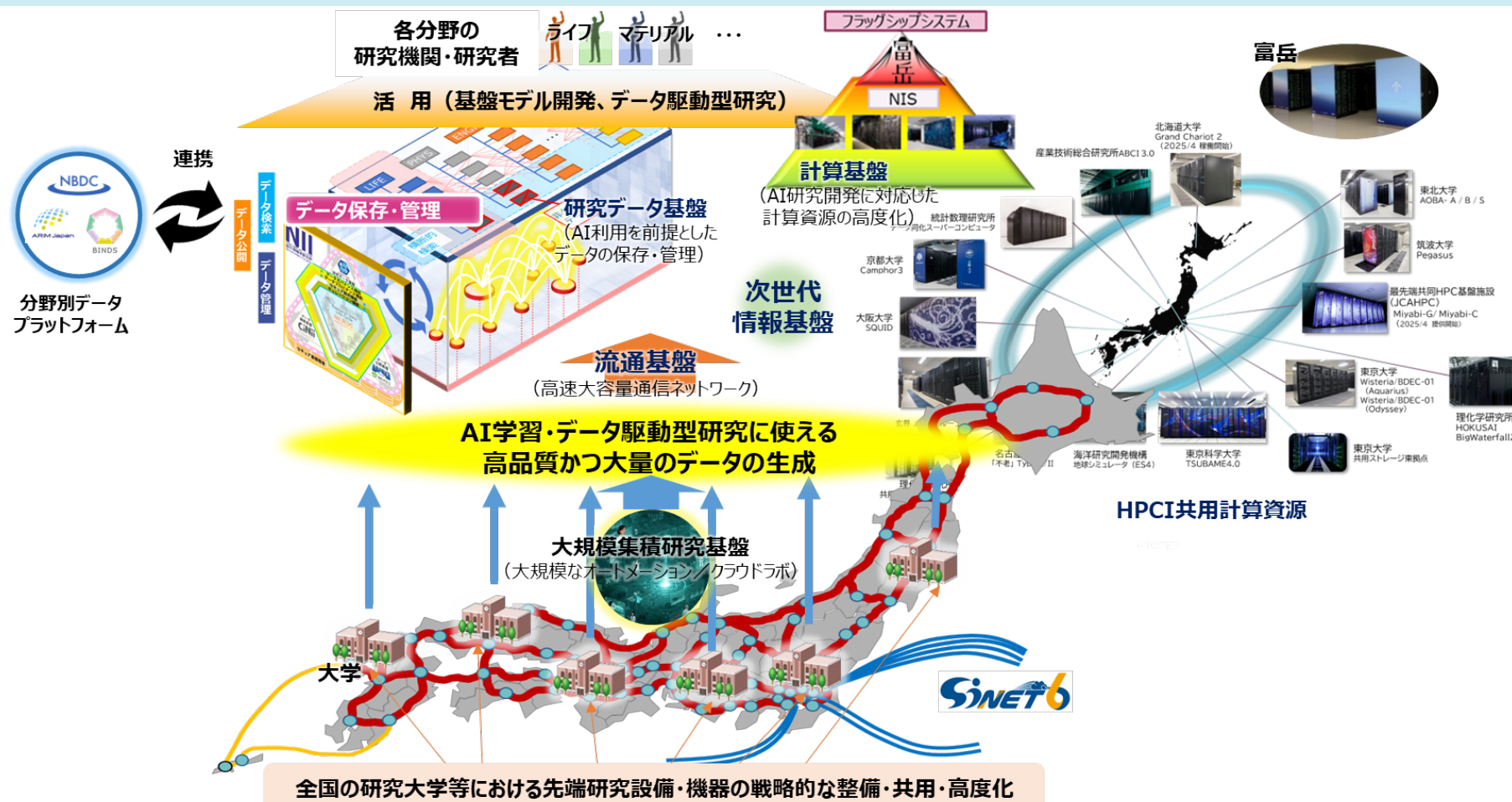


※文部科学省におけるAI for Science推進の動きとも適時連動

以下、参考

AI for Science を支える次世代情報基盤の構築：流通基盤の強化

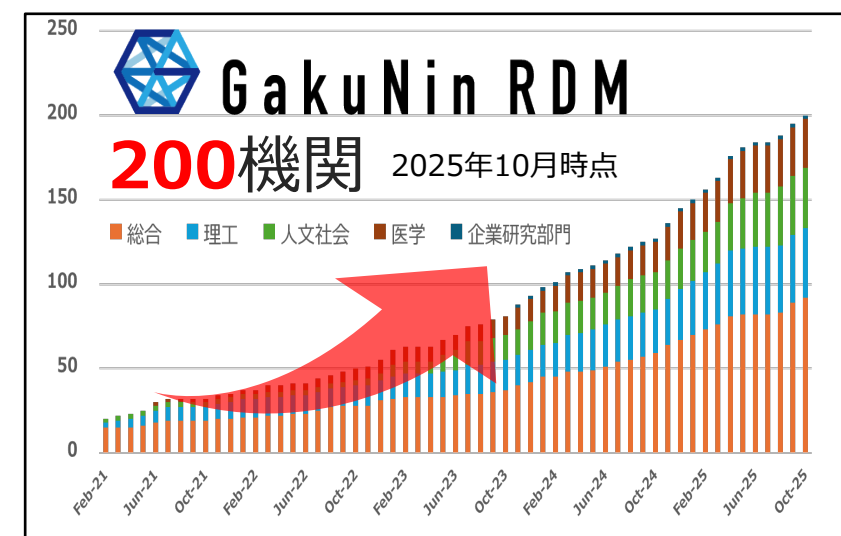
- 各地の研究大学等及び大規模集積拠点における最先端の研究設備・機器から創出される**高品質かつ大量のデータ**を、全国に張り巡らされた**流通基盤**を通して**研究データ基盤**に蓄積し、**計算基盤**等によりデータの利活用を促進していくため、AI for Scienceを支える次世代研究インフラの構築は不可欠。
- **AI for Science を支える研究データの管理・利活用と流通の在り方について早急に検討**を行い、全国の研究者等にとって簡便に使いこなすことができる**AI 時代に適した研究データ基盤 NII RDC** や**流通基盤 SINET** の高度化を推進する。



AI for Science を支える次世代情報基盤の構築：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AI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**。海外のAI戦略においても、高度なAIモデルの開発等のために、AI対応可能なデータセットの構築の必要性が明記。
- 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AI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**。また、研究者の安定的なAI・データ駆動型研究と、オープンサイエンスの推進を担保し、我が国のデータ主権を確保することは**経済安全保障上も極めて重要**。
- そのため、我が国の研究データの管理・利活用のための中核的なプラットフォーム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**AI対応可能な研究データを保存・管理する仕組み（メタデータ付与）の導入等**を進めていく。
- AI for Science の推進に向けたNII RDC の高度化により、国内外の研究機関・大学や産業界の**知の結節点**とし、AI時代の**国際共同研究や産学連携を促進**する。

海外の研究機関・大学



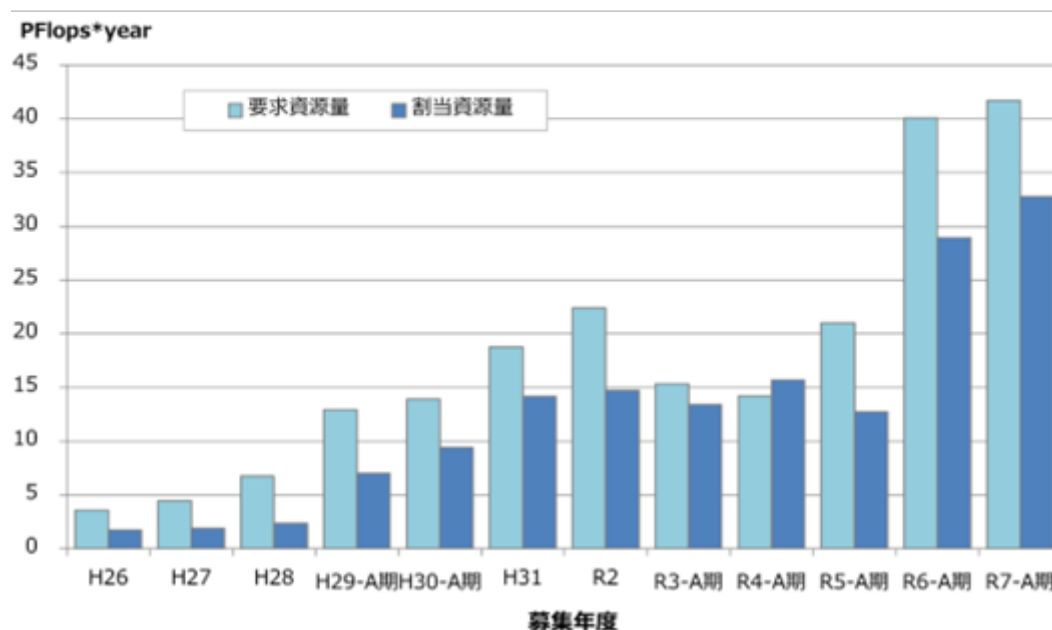
AI for Science を支える次世代情報基盤の構築：計算資源の戦略的増強

- 科学基盤モデルの開発・利用等の研究活動におけるAI利活用（AI for Science）を本格的に進めるためには、**AI向けの膨大な計算資源を有する計算基盤が必要不可欠**。
- 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（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）の利用状況は既に逼迫しており、AI for Scienceに向けた**計算資源の戦略的な増強及び効率的な配分**が喫緊の課題。
- AI for Scienceの実現に向けた研究開発の取組に必要となる計算資源等の確保に向けて、短期実装と中長期的な全体底上げを見据えつつ、**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ことが必要である**。

HPCI共用計算資源



HPCIの要求資源量と割当資源量の推移



(参考) AI for Science に関する有識者ヒアリングの概要

AI for Science を支える次世代情報基盤の構築

- AI for Scienceに関して、情報分野の研究者や AI を用いた科学研究に取り組んでいるドメイン研究者等、産学の幅広い専門家約70名以上からヒアリング実施。
- このうち、次世代情報基盤の構築に関連する内容を一部抜粋

期間	: 2025年8月下旬～10月上旬
ヒアリング先	: 情報分野の研究者や、AI を用いた科学研究に取り組んでいるドメイン研究者（ライフ分野、マテリアル分野、数理分野、人文社会系、等）、産業界 等（約70名）
ヒアリング内容	: AI for Scienceに関する現状・将来像、課題や問題意識、必要な支援等について

データ基盤

- ・ 生み出された研究データを国内で保持するために、国全体として研究データ基盤の強化が必要。
- ・ 研究データの蓄積・活用→新たなデータの創出・蓄積、という循環の構築に期待。データ作成者・提供者が研究データを公開しようというモチベーションにつながり、高価値のデータ活用が進むのではないかな。

基盤の連携

- ・ 日本はデータ、AI、ロボット、実験装置、計算資源がつながっていない。
- ・ 今後、データの探索と実験の両方でAIを使おうという動きが出てくるので、データ基盤と計算基盤の連携が求められるのではないかな。
- ・ 基盤・裾野部分を広くすること、富岳NEXTのような先端を引き上げることの両方の視点が必要。

計算基盤

- ・ 日本国内では計算資源（GPU、メモリ）が不足しており、クラウドの計算資源を柔軟に確保できるサポート体制・仕組みが求められている。
- ・ 富岳 NEXTにはAI for Scienceに用いることができる大規模な計算資源として期待している。
- ・ 多くの研究者が HPCI へのアクセスに障壁を感じており、クラウドサービスや個人で購入した PC 等を使っている状況。
- ・ 理論研究やAIの利用においては大きな計算資源は必要としないことが多い。
- ・ 予算制度がクラウド利用の足枷になっているのではないかな。複数年度にわたるクラウド契約には基金化が必要である。

<データ戦略>

- AI ロボット開発のボトルネックはデータであり、デジタルツインを活用してデータの収集を進めることが必要。
- データをやみくもに取得するのではなく、必要なものを精査して賢く集めることが必要。
- これまでに蓄積してきた独自のデータをいかに活用するかという点に重点を置くことが重要。
- 現状、AI が直接利用可能な形式で整備されているデータがほばない。
- 国際的な潮流として、研究成果を構造化されたデータとして蓄積・活用する動きが進んでおり、日本もこの流れに乗り遅れないよう、研究データの整備と利活用の推進、データ戦略の検討・策定を行うべきである。
- ライフサイエンス分野において、厳格な個人情報保護法や倫理委員会の規定が AI 活用の障壁となっている。
- これまで進めてきた研究データの管理・利活用やオープンサイエンスはAI時代のデータの在り方に合った流れである。
- データ同士、データと他の知見との掛け合わせ等の有機的な結合により日本としてユニークで競争力あることができるのではないかと。
- AIや計算機の進展により、これまで取れていなかったデータが取れるようになってきた。取得したデータを死蔵させることのないように、データ管理することにもっと研究費を出していけるようにすべき。
- 人文社会系については論文や資料のデータ化が遅れており、著作権の関係もありAI活用の障壁となっている。日本語信頼性向上には豊富に存在する人文学資料のデジタル化とデータ整備が急務。
- データベースの構築の際には、ドメイン研究者だけでなくAI や統計の専門家、他分野の人も入ってオープンアンドクローズ戦略を議論することが重要。
- データベースそのものも研究要素であり、将来有効に使えるデータセットを作ることがAI研究と同じくらい重要。
- 企業はデータの外部提供に慎重。連合学習等で秘匿性を担保しつつ、失敗例を含むネガティブデータの価値を明示し、コンソーシアム参加などのインセンティブ設計でデータ提供を促すとよいのではないかと。
- 一部の学問領域については、非競争領域として民間のデータをうまく取りこむことが可能であり、領域全体のデータ戦略が重要。
- プロジェクトが終わったときに構築したデータセットを国としてどう維持・活用していくかが重要な戦略になる。
- 研究データ管理にもAIの活用を進めるべき。

経済財政運営と改革の基本方針2025（令和7年6月13日閣議決定）

第2章 第2章 賃上げを起点とした成長型経済の実現

3. 「投資立国」及び「資産運用立国」による将来の賃金・所得の増加

（4）先端科学技術の推進

（前略）イノベーションの持続的な創出に向け、国際卓越研究大学制度による世界最高水準の研究大学の創出を始め多様で厚みある研究大学群の形成に向けた取組を、効果検証しつつ進めるとともに、先端研究設備・機器の戦略的な整備・共用・高度化を推進する仕組みを構築する。研究データの活用を支える情報基盤の強化やA I for Scienceを通じ、科学研究を革新する。産学官連携の大規模化・グローバル化を促進する。

科学技術人材の育成を強化する。成長分野における大学学部・高専学科の再編及び高専の新設、先端技術に対応した人材育成【注1】の高度化・国際化を始め、大学・高専・専門学校の機能を強化する。

【注1】統計人材の育成を含む。

新しい資本主義のグランドデザイン及び実行計画2025改訂版（令和7年6月13日閣議決定）

Ⅲ. 投資立国の実現

3. G X・D Xの着実な推進（2）D X ①A I イノベーション促進とリスク対応の両立

ii）計算資源・情報通信基盤等の整備

研究データ基盤や計算基盤等の施設・設備等の整備や共用、ワット・ビット連携、データセンター等の整備を加速する。

質の高い日本語データの整備・拡充や未利用データの活用等に加え、日本の文化・習慣等を踏まえた信頼できるA I 開発・評価の推進・活用を進める。

V. 科学技術・イノベーション力の強化

3. 大学等の高度な研究・教育と戦略的投資の好循環の実現

①大学ファンドによる支援と地域中核・特色ある研究大学への支援

世界最高水準の研究大学の実現に向けて、10兆円規模の大学ファンドの支援対象となる国際卓越研究大学の第2期公募における選定を進め、2025年度中の助成開始を目指すとともに、意欲ある多様な大学による、各々の強みや特色を十分に発揮し、地域の経済社会の発展や国内外における課題の解決や研究の多様な国際展開を後押しする。加えて、研究大学や大学共同利用機関法人（個々の大学では整備できない大規模施設・設備等を全国の研究者に提供する機関）等における先端研究設備・機器の戦略的な整備・共用・高度化を進めるとともに、技術専門人材の育成・情報基盤の強化やA I for Scienceを通じ、科学研究を革新する。

統合イノベーション戦略2025（令和7年6月6日閣議決定）

2. 第6期基本計画の総仕上げとしての取組の加速

（1）先端科学技術の戦略的な推進 ② 研究施設・設備の強化、オープンサイエンスの推進

（学術論文等のオープンアクセス化の推進）

- 「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）」に基づき、令和7年度新規公募分からの学術論文等の即時オープンアクセス実施に向けて、学術プラットフォームに対する大学等主体の拡大した集団交渉体制の構築支援、機関リポジトリ等の情報基盤への論文掲載やシステム間連携の実装、研究成果発信プラットフォームの整備・充実を進める。

（公的資金による研究データの管理・利活用の推進）

- 先行的な取組であるムーンショット型研究開発制度における先進的データマネジメントの実施状況を検証するとともに、S I P第3期におけるデータマネジメントの推進、公的資金により得られた研究データへのメタデータ付与、大学等の研究開発機関におけるデータポリシー策定と機関リポジトリへの研究データ収載、G7等の国際連携等により、研究データの管理・利活用を推進する。

（2）知の基盤（研究力）と人材育成の強化 ② 研究施設・設備の強化、オープンサイエンスの推進

（研究DXを支えるインフラ整備や研究施設・設備の共用化の推進）

- A I・データ駆動型研究による効率化・迅速化を推進するため、超高速・大容量のネットワーク基盤（S I N E T）や全国的な研究データ基盤（NII Research Data Cloud）といった研究デジタルインフラの高度化等を進める。
- 「富岳」を効率的かつ着実に運用し、幅広い活用を促進するとともに、優れたA I性能を有する次世代フラッグシップシステムの開発・整備を進める。また、運用開始後の成果創出を見据えたアプリケーション開発支援、人材育成等を推進する。

3. 第7期基本計画に向けた議論も踏まえた取組の推進

（2）研究力の強化、人材の育成・確保 ① 大学等の運営・研究基盤の強化

- 研究データの適切な保存・管理、流通、活用を促進する情報基盤等の強化を進めるとともに、ライフサイエンス・マテリアル等の分野を含む研究データを活用した科学研究向けA I基盤モデルの開発・共用等のA I for Science を加速させ、科学研究の革新につなげていく。

科学の再興に向けて 提言（令和7年11月18日「科学の再興」に関する有識者会議）

（４）時代に即した研究環境の構築

（４）－１ AI for Science による科学研究の革新

③ AI for Science を支える次世代情報基盤の構築

科学基盤モデルの開発・利用等の研究活動における AI 利活用を本格的に進めるためには、AI 向けの膨大な計算資源を有する計算基盤や、増大する研究データの保存・管理を行う研究データ基盤、それらの流通を支える大容量の情報通信ネットワークを提供する流通基盤が必要不可欠である。

我が国には、研究データの管理・利活用のための中核的なプラットフォームの研究データ基盤（NII RDC）や、日本全国の大学・研究機関等を超高速・低遅延でつなぎ、流通させるSINET、世界最高水準のスーパーコンピュータ「富岳」が存在する。AI for Science の実現に向けて、これらの情報基盤を効果的に活用するとともに、早期の取組の加速と中長期的な全体的な底上げを見据え、HPCI（ハイパフォーマンス・コンピューティングインフラ）を中心とした共用計算資源の戦略的な増強等、ユーザビリティの高い計算資源の利活用を推進するとともに、「富岳」の次世代となる新たなフラッグシップシステムの開発・運用を通じた AI 処理能力・アプリケーション実効性能の飛躍と国産技術の国際市場への訴求を図る。

また、AI 時代においては、研究システムの自動・自律・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が想定されるとともに、AI の高度化に向けては、今まで以上に高品質な研究データの管理・利活用が求められるところ、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を行い、AI 時代に適した研究データ基盤 NII RDC や流通基盤 SINET の高度化を推進する。

（４）－２ 研究施設・設備、研究資金等の改革

研究体制を十分に整えることが難しい若手研究者や自ら設備投資・運用コストを負担することが必要な研究者にとって、コアファシリティによる支援は極めて重要であり、欧米や中国に対して日本の研究環境の不十分さが指摘される要因となっている。このため、第7期基本計画期間中に、我が国の研究基盤を刷新し、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、技術職員や URA 等の人材を含めたコアファシリティを戦略的に整備する。その際、取組が効果を最大限発揮するためには、組織としての経営戦略の構築・実装や研究資金の改革と併せて一体的に対処することが肝要となる。あわせて、研究活動を支える研究設備等の海外依存や開発・導入の遅れが指摘される中、研究インフラのエコシステム形成に向けて、産業界や学会、資金配分機関（FA）等とも協働し、先端研究設備等の整備・利活用・高度化・開発を推進する。AI for Science が急速に進展する中、高品質な研究データを創出・活用するため、全国の研究者の研究設備・情報基盤等へのアクセスの確保や計測・分析等の基盤技術の維持は、経済安全保障上も重要である。