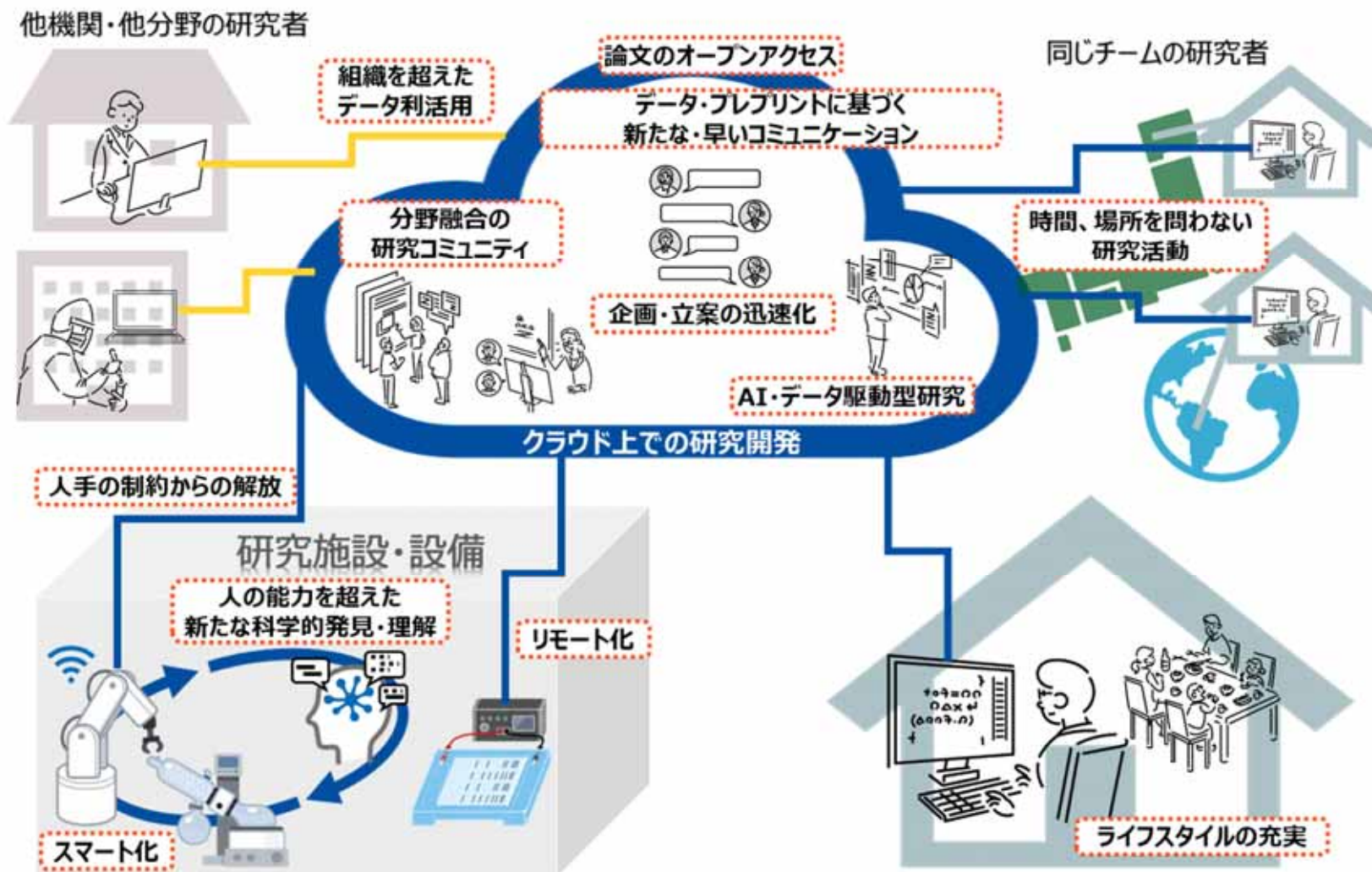


研究DXの推進について

令和4年6月

研究デジタルトランスフォーメーション（研究DX）の考え方

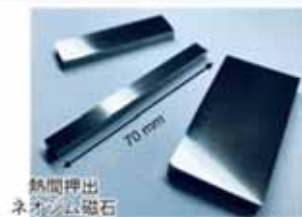
「AI」×「データ」×「リモート化・スマート化」⇒ 価値創造
研究デジタルトランスフォーメーション（研究DX）



研究DXの先行的な成果事例

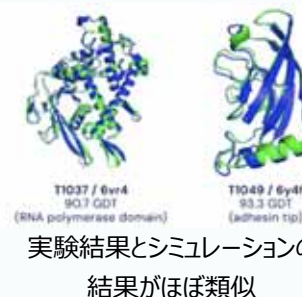
AIが約6600万通りから最適な作製条件を探索 → AI解析前後で材料性能が1.5倍に向上

- 世界の電力消費量の約50%を占めるモーターの効率化・省エネ化のために、永久磁石の高性能化が期待。
- ネオジム磁石の作製プロセス条件は約6600万通りであり、網羅的な探索が不可能。AI解析により提案された最適な作製条件に基づいて実験計画を立てることで、わずか40回程度の実験でネオジム磁石の強さをAI解析前の約1.5倍向上させることに成功。
(2021.11.15 プレスリリース)



AIによるタンパク質の立体構造解析ツール → 数年かかっていた作業を圧倒的に短縮

- 2014年にGoogleが約670億円で買収したAI開発企業DeepMindが無償公開した遺伝子配列情報からタンパク質の立体構造を解析するAI「AlphaFold2」。
- 従来、タンパク質のアミノ酸配列の構造を特定するには数か月～数年かかり、多大な時間と費用が掛かっていたが、AlphaFold2では、限られた情報から構造と機能を推定することが可能に。



スパコン「富岳」を活用したゲリラ豪雨予測 → 30秒ごとに更新するリアルタイム性を実現

- ゲリラ豪雨は5～10分といった短時間で状況が急激に変化するため予測が困難。理研等の研究グループは、マルチパラメータ・フェーズドレイ气象レーダによる雨雲の詳細な観測データと、「富岳」による大規模計算により、高頻度かつ高精度な予測を実現。
- 2021年夏に、首都圏において30秒ごとに更新する30分後までのリアルタイム降水予報を世界で初めて実施。
(2021.7.13 プレスリリース)



実証実験で表示される「3D 雨雲ウォッチ」アプリイメージ

仮説・実験・検証・修正のサイクルを自動化 → これまで着手が困難だった希少疾患等の創薬開発が推進

- ケンブリッジ大学とマンチェスター大学により開発されたロボットサイエンティスト「イブ」は、一日当たり1万化合物のスクリーニングを行うとともに、実験結果を統計学・機械学習により解析し、より高い活性をもつ新たな化合物の構造を予測し、実験・検証・修正のサイクルを繰り返す、仮説主導のハイスループット研究を自動化した。
- このような手法を導入することにより、製薬企業が手を出しにくい希少疾患等に対する創薬開発が推進されることが期待。



Robot scientist Eve

「AI」、「データ」、「リモート化・スマート化」による圧倒的な生産性の向上とハイインパクトな成果が生まれ始めている。しかし、まだ一部の研究者・研究領域のみ。この変革の動きを、日本全体に発展させることが必要。

研究DXの課題と必要なアプローチ

- AI技術の活用や研究データの共有など研究DXの取組を進めるに当たっては、具体的に何をすればよいのかわからない、メリットが不明確である、研究データの性質に応じた公開・非公開のバランス、データ形式の共通化にかかるコストなど様々な課題が存在。
- これらを解決し研究DXを全国的な動きにするためには、「ユースケースの形成、普及」、「データ共有・利活用の促進」、「研究デジタルインフラ等の効果的活用」を一体的に進めることが必要。

【研究DXを進める上での主な課題】

研究DXに対する理解が浸透していない

そもそも研究DXとは何か、何をすればよいのか、といった研究DXに対する認識や理解が浸透していない。

研究DXのメリットが不明確

研究DXはまだ萌芽的なフェーズであり、具体的なケースが乏しく、そのメリットや意義が理解されにくい。

オープンアンドクローズの難しさ

分野やデータ特性によって、公開・非公開の戦略が求められるが、その考え方やノウハウが蓄積されておらず、効果的な研究データの共有や利活用が進みにくい。

研究DXの実施にかかるコスト

メタデータの付与やデータの構造化など、研究DXに必要な人手や時間といった追加的負担をどのようにカバーするか。

計算資源やストレージの確保

研究DXが進めば進むほど、取り扱うデータの量が増加し、必要となる計算資源やストレージの規模が大きくなってしまふ。

分野と基盤（情報、計測等）の研究者の連携

各分野における研究DXを進めるためには、分野の研究とAI等の情報科学、計測等の実験装置開発の連携が不可欠だが進みにくい。

【課題解決に必要なアプローチ】

1. 価値創造を目指したユースケースの形成、普及

- 具体的な研究DXの成功事例（産業界の参画、参加者のインセンティブ形成、海外連携含む。）を創出し、その成果や過程で得られる仕組みや経験・ノウハウを普及することが必要ではないか。

2. データ共有・利活用を促進する基盤的機能の強化

- 研究データの共有・利活用に取り組むよう研究者や大学、研究機関の活動を促すとともに、それらを支援する機能の整備や共通的な課題への対応が必要ではないか。
- また、分野・基盤の研究コミュニティの融合による革新的な研究手法の開発を促進することが必要ではないか。

3. 研究デジタルインフラ等の効果的活用

- 我が国には先端的な研究インフラがすでに数多く整備されている。これらを効果的に活用する方策を進めることが必要ではないか。

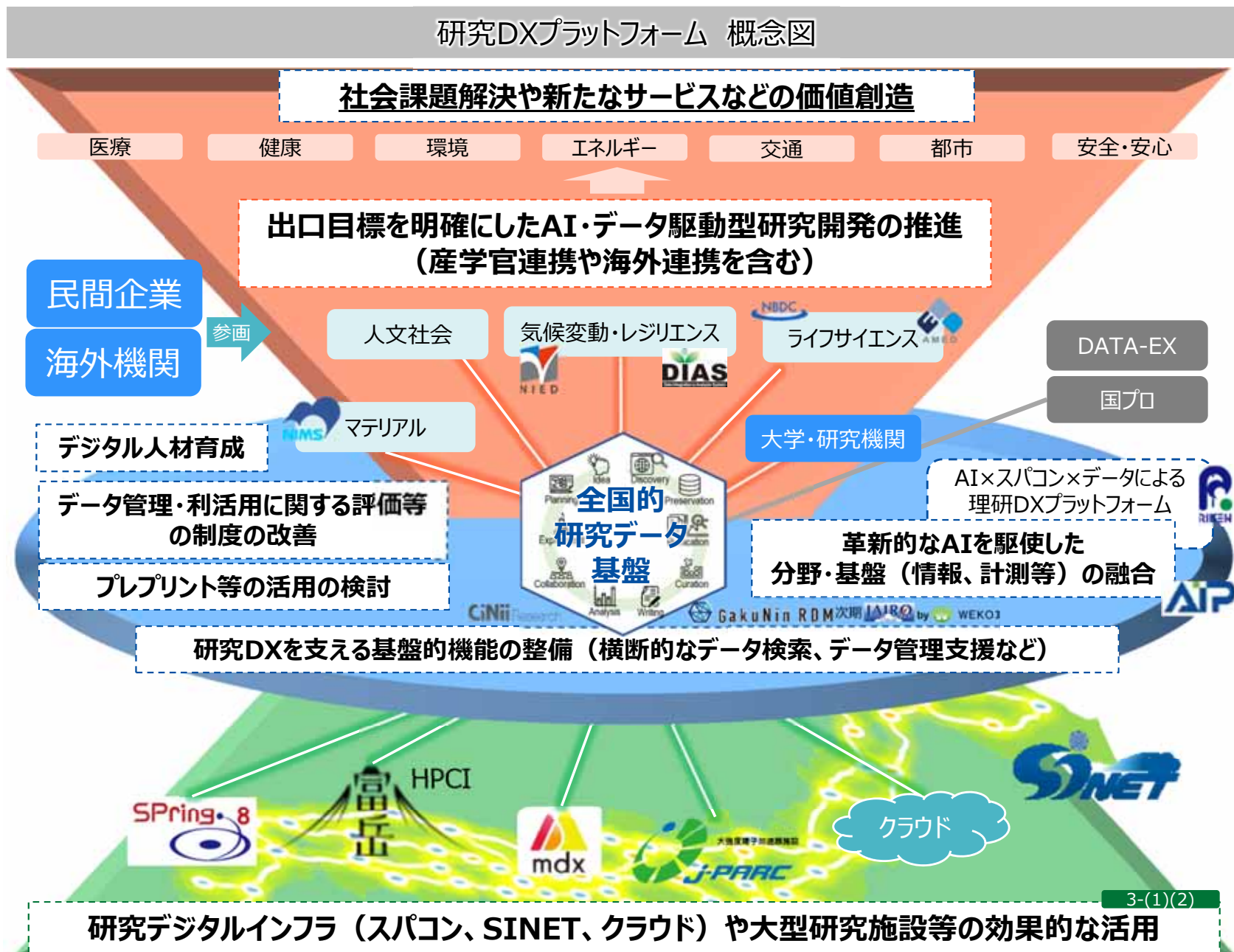
研究DX実現に向けた取組 概念図

研究DX実現に向けたアプローチ

1. 価値創造を目指したユースケースの形成、普及

2. データ共有・利活用を促進する基盤的機能の強化

3. 研究デジタルインフラ等の効果的活用



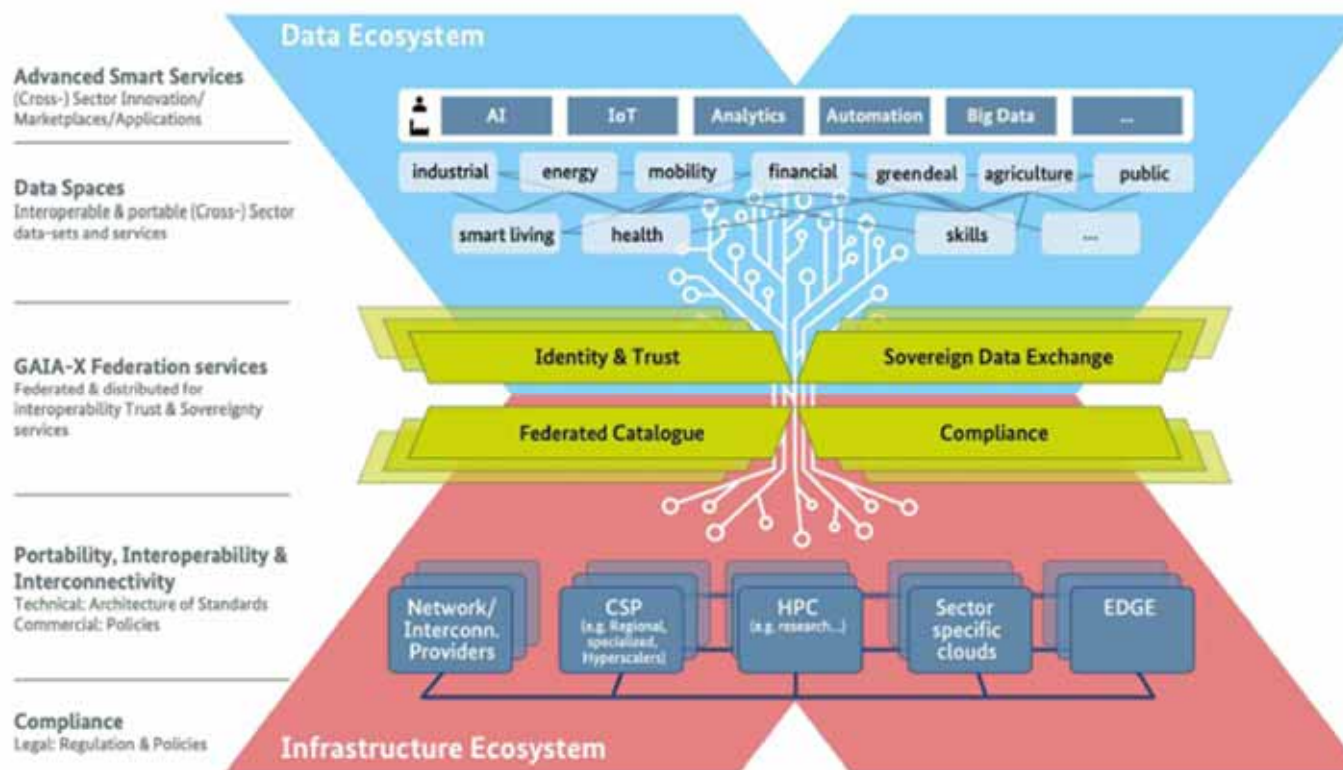
※ 上記の取組の進捗をフォローアップし、研究DXの進展による研究スタイルの変容を踏まえた、研究評価や研究プログラムの在り方について、随時検討を実施

参考資料

【参考】GAIA-X概要

<概要>

- 欧州域内外の企業のさまざまなクラウドサービスを単一のシステム上で統合し、業界をまたがるデータ交換を容易に行える標準的な認証の仕組みを通じて、相互運用性を実現するための基盤
- 各産業部門から生成されるデータの相互運用等を実現する「データエコシステム」レイヤー、エッジコンピューティング等の相互運用を実現する「インフラエコシステム」レイヤー、データインフラを利用する際のセキュリティ等を定める「フェデレーションサービス」レイヤーで構成



GAIA-Xについて（DSA関係者より聴取）

- Industry4.0で産業データ連携のコネクター開発をしてきたが、データの対象範囲をスマートシティなどに拡大した。
- 欧州はGAFAの存在も意識し、クラウドに置かれたデータのコントロール権を持つというGAIA-Xのコンセプトが生まれた。
- ビジネスドメイン（例えば自動車会社）毎のコミュニティでポリシー・ルールといった議論を行っており、GAIA-Xという1個のシステムがある訳ではない。
- データ流通のための技術的な検討や標準化は別の機関で実施。

気候変動対策・レジリエンスに貢献するデータ解析プラットフォームの形成

【概要】

気候変動やそれに伴う極端気象の激甚化・広域化、及び地震・津波・火山等の自然災害への対応が必要であるが、研究開発・意思決定・ファイナンスに必要なインテリジェンスの不足が大きな課題となっている。

レジリエントで安全・安心な社会やカーボンニュートラルの実現に向け、研究開発・実務現場での地球環境ビッグデータ利活用の促進やインテリジェンスの創出を目指し、新たな技術・価値を生み出すデータ・解析プラットフォームを形成する。



マテリアルDXプラットフォームの構築

- ✓ 我が国が世界に誇る計算基盤や研究データベース、先端共用施設群や大型研究施設等のポテンシャルと強みを相乗的に活かし、世界を先導する価値創造の核となる「**研究DXプラットフォーム**」を構築
- ✓ 幅広い課題解決に貢献するマテリアル分野をユースケースとし、①データ創出から、②データ統合・管理、③データ利活用まで、一貫通貫した研究のDXを推進

➢ 富岳、mdx等のポテンシャルを最大限活用し、社会課題解決や未開拓分野・融合分野のデータ駆動型科学を推進



研究ポテンシャル・強みを掛け合わせるデータ統合プラットフォーム

全国的な研究データ基盤を構築・高度化
➢ 分野別リポジトリや計算資源との連携
➢ 各種コンプライアンス対応

先端大型共用施設のポテンシャル最大化・DX基盤強化



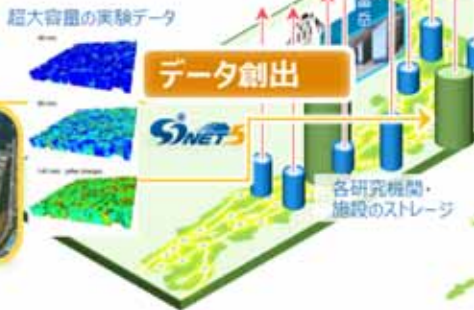
データ利活用

未来の価値創造を先導するデータ駆動型研究開発の推進

データ統合・管理

データ検索
データ公開
データ管理

データ創出



トポサイエンス
トポイノベーション
の価値的創出

カーボンニュートラルなど将来の社会課題解決
➢ 計算・計測・データサイエンスが融合する新たな研究の方法論により未来価値創造を先導



未知の構造・関係の可視化

マテリアルDXプラットフォーム
マテリアル分野のAI・データ駆動型研究による研究手法の革新

ハブ&スポーク

➢ 全国の先端共用設備から創出されるデータを解析可能な形で全国共有

ライフサイエンス分野におけるユースケースの形成・普及

【概要】

有用な研究データの幅広い利活用の推進を目指し、AMEDを中心として府省横断的に進められているヒトゲノム情報等のデータ利活用に関する取組を推進するとともに、AIとビックデータ、実験装置等の自動化等を組み合わせた革新的な研究アプローチのユースケース創出を推進する。

研究データ管理・統合
データ利活用の推進

健康・医療情報 ※個人情報保護に留意

- ヒトゲノム情報や臨床情報等のデータ利活用に関する取組を推進（内閣府とりまとめ） <Visiting型計算環境、検索機能、カタログ化、ワンストップ窓口>

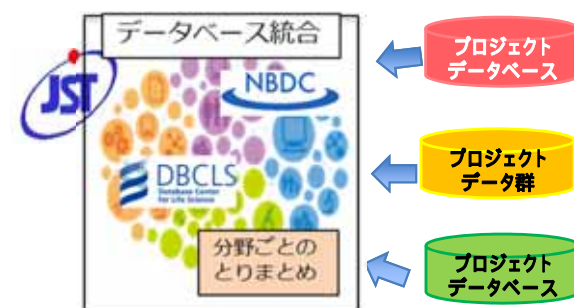


- 健康情報等の集約・匿名化（次世代医療基盤法（内閣府とりまとめ）活用等）

文部科学省支援

その他ライフサイエンス関係データ

- 「ライフサイエンスデータベース統合推進事業」において、効率よく研究者、産業界、さらには国民に還元していくための統合的なデータベースの構築・利活用促進
- 研究データ基盤との連携検討（データ検索機能等）



- 戦略目標「バイオDX」領域の推進とモデルとして活用
- 理研において、メタデータを付与したデータ（脳画像等）や解析ツールを公開



多様な研究データを収集するとともに、ユーザーが使いやすい高品質な研究データを提供

ユーザーニーズを踏まえた更なるデータの創出

研究データ
創出



クライオ電子顕微鏡

- ロボットや遠隔技術による実験の自動化
 - 計測機器・技術の高度化による高解像度・時系列のデータの大量取得
- 【取組例】
- ✓ クライオ電子顕微鏡における試料調製の自動化、操作の遠隔化
 - ✓ 理研が有する研究基盤・装置群と連携し、創薬、ワクチン関連研究や、バイオものづくりに資する分子デザイン・化学合成を推進



理研の強みを特に生かせる領域

- 多階層（分子～個体）・マルチモダル（オミクス、イメージング等）データの高効率計測・統合解析によるブレイクスルー

<毛包の形成過程>
1細胞遺伝子発現解析と時系列イメージングの統合による高解像度4Dイメージング



<ロボット実験>
ロボット技術と画像AIの統合による高効率・高品質実験



分子デザインをユースケースとする理研DXプラットフォームの構築

【概要】

「AI×スパコン×研究データ」による革新的な成果創出を目指し、理研DXプラットフォームの構築を推進するとともに、創薬・ワクチンや、バイオものづくり等に資する分子デザイン・化学合成における研究DXのユースケース創出を推進する。

