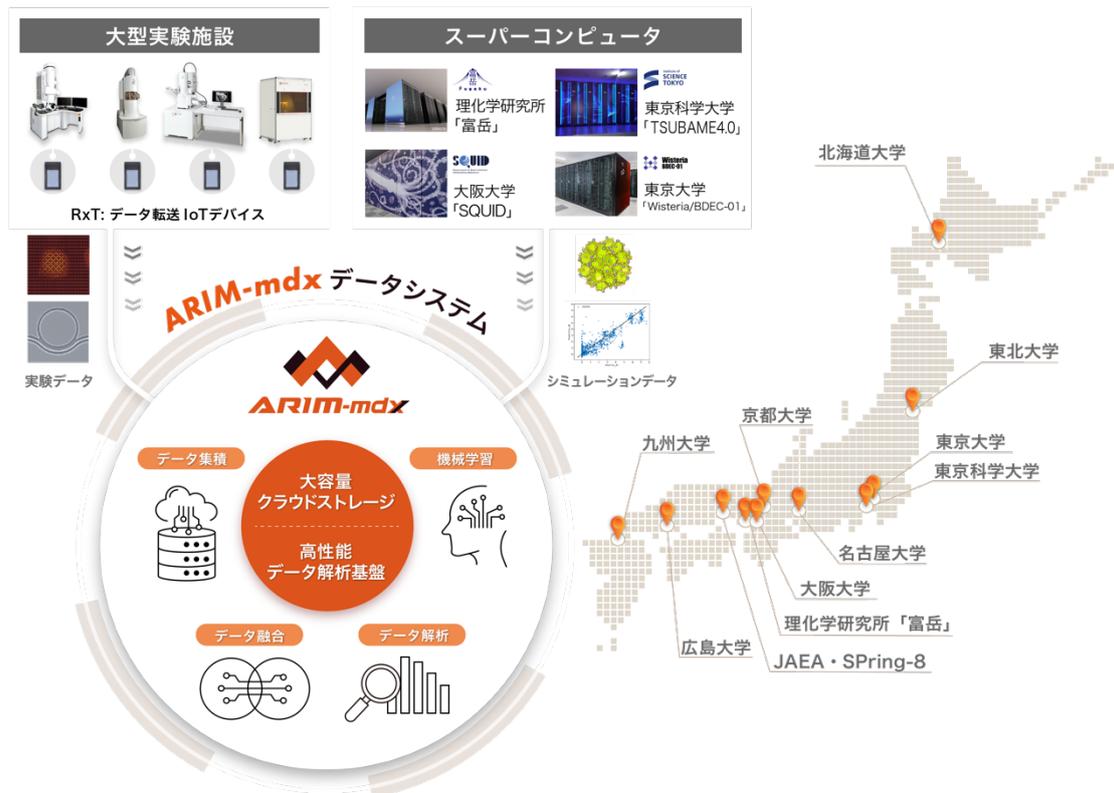


材料研究 DX を加速する ARIM-mdx データシステムを開発、 全国の 900 名以上が利用開始

発表のポイント

- ◆全国の材料科学の研究者が大型実験装置、スーパーコンピュータ、学術クラウドを連携して使える先駆的材料研究用大規模データプラットフォーム「ARIM-mdx Data System」を開発し、約1年にわたる運用で、産学140以上の機関・企業から900名以上の研究者が利用した。
- ◆「データ活用社会創成プラットフォーム mdx」の上に構築され、実験装置からのデータ転送を自動化するIoTデバイスの導入や、全国の主要スーパーコンピュータとの学術情報ネットワーク SINET6 による接続により、研究現場の効率を大幅に改善する。
- ◆データ駆動型の材料研究や、AI・機械学習を活用した材料探索、実験と理論の密接な連携による新しい研究手法の確立など、材料研究のデジタル・トランスフォーメーション (DX) の加速につながると期待される。



ARIM-mdx Data System 概要

概要

東京大学情報基盤センター、同大学大学院工学系研究科、日本原子力研究開発機構、広島大学、理化学研究所、情報・システム研究機構 統計数理研究所のメンバーからなる研究グループは、材料科学分野の研究 DX を加速する全国規模のデータプラットフォーム「ARIM-mdx Data System」を開発しました。2023年8月の運用開始以降、約1年間で産学140以上の機関・企業から900名以上の研究者の利用がありました。

本システムは、文部科学省のプロジェクト「マテリアル先端リサーチインフラ (ARIM Japan) ^{注1}」の一環として整備され、高性能な計算機と大容量のストレージを備え、SINET6の高速ネットワークと連携してデータを収集できる「データ活用社会創成プラットフォーム mdx ^{注2}」の上に構築されました。全国の大学・研究機関の実験施設とスーパーコンピュータをつなぎ、研究データの利活用を促進します。

本研究成果はビッグデータに関する国際会議 IEEE BigData 2024(2024年12月15日-18日、米国ワシントン DC) にて発表されます。

発表内容

【研究の背景・取り組むべき課題】

材料研究を取り巻く環境の変化

近年、材料科学分野では研究データの重要性が飛躍的に高まっています。画像認識や自然言語処理などの分野で機械学習が大きな成功を収めている背景には、大量のデータの存在が不可欠でした。この潮流は材料研究分野にも波及しており、大量の実験データと計算データを効率的に収集・利活用できる包括的なデータシステムの必要性が増大しています。

材料研究における実験的手法と理論的手法の連携

現代の材料研究手法は高度化しており、実験的手法と理論的手法を組み合わせることが不可欠です。実験的手法では、電子顕微鏡や大型放射光施設などの最先端実験装置を駆使し、実際の材料の構造や特性を直接観察・測定します。また、高度な加工装置によってデバイスを作製します。理論的手法では、スーパーコンピュータを活用し、大規模な計算を行うことで材料の複雑な構造予測や物性・プロセス条件の分析をおこないます。実験的アプローチと理論的アプローチを組み合わせることで、背後に隠れた物理現象の理解や新材料の開発を促します。

研究データ管理における課題

これまで、材料研究における科学データの収集は個別の分野や研究グループで独立には進んでいたものの、幅広い材料研究者が大規模データを収集・管理可能なデータプラットフォームは未整備でした。その結果、多くの研究グループにおいて、データは個々の実験施設やスーパーコンピュータ、個人のPCに分散して保存され、研究メンバーの増加やデータ量の肥大化に伴うデータ管理のコストが問題となっていました。

【ARIM-mdx Data System】

これらの課題を解決するため、研究グループは ARIM-mdx Data System を開発しました。その特徴は以下のとおりです。

実験データと計算データを一元管理、フルリモートで解析

- 全国の大型実験施設とスーパーコンピュータからの研究データを ARIM-mdx Data System に集約し一元管理
- 実験施設 ↔ スーパーコンピュータ ↔ ARIM-mdx Data System は学術ネットワーク SINET6 で接続され、超高速データ転送を実現
- ブラウザからワンクリックで高性能データ解析サービスを立ち上げ、世界中どこからでも収集されたデータをフルリモートで解析可能

実験データの収集・共有の効率化

- 研究施設の実験装置から IoT デバイスを使って、データを安全かつ自動的に転送・保存
- 実験中の即時データ解析を実現

実験・計算の分野間や産学間の垣根を越えた研究基盤

- 実験・計算分野の研究者に共通のプラットフォーム
- 大学・企業の研究者が同じプラットフォームでデータを活用
- 組織や分野の枠を超えた共同研究を促進

【これまでの実績】

ARIM-mdx Data System を開発、運用開始することで、データ駆動型の材料研究における新たな研究スタイルを確立しました。これまでに、東京大学内の 40 台以上の共用実験装置を ARIM-mdx Data System に接続したり、国内 9 つの HPCI^{注3} / JHPCN^{注4} スーパーコンピュータからの大容量データ転送を実現したりするなどの実績があり、約 1 年の運用で産学 140 以上の機関・企業から 900 名以上の研究者が利用しました。

【研究者コメント】

本研究論文の主著者、東京大学情報基盤センターの華井雅俊 特任助教は本研究成果について「コンピューターサイエンス分野と材料科学分野の様々なメンバーが大学や組織の垣根を超えて協力しサービスを実現することができました。今日の機械学習や AI の盛り上がりは材料科学分野にも波及し先進的な研究が進められる一方で、データ収集のリモート化などといったもっと基本的な部分に課題が多くありました。本システムの利用拡大により、新材料開発の効率化や、データ駆動型の材料研究が促進されることを期待しています。特に、AI・機械学習を活用した材料探索や、実験と理論の密接な連携による新しい研究手法の確立を支援することを目指していきたいと思います。今後、北海道大学、電気通信大学、産業技術総合研究所、名古屋大学、広島大学、日本原子力研究開発機構 SPring-8、九州大学の大型実験施設などへの導入拡大を予定しており、本サービスを全国の材料研究者の方々に使っていただければ開発者として非常に嬉しく思います。」と述べています。

本論文の共著者で ARIM 事業の東京大学 データ基盤部門の統括責任者、東京大学大学院情報理工学系研究科（情報基盤センターを兼務）の田浦健次郎 教授は「ARIM-mdx システムは材料科学・工学の研究者と、情報科学の研究者の密接な協力によって生まれました。実際に装置を使う研究者と議論をし、その目線に立って、実験・観測データの取得から ARIM-mdx システムへのアップロード作業を自動化する IoT デバイスを開発しました。そのための基盤として学術機関向けのクラウドである mdx を使い、開発を迅速に行うとともに、AI・機械学習・シミュレー

ションを用いた材料探索・材料科学の研究に広く貢献して行くことを期待します。」とコメントしています。

ARIM 事業の東京大学代表者であり論文共著者の、東京大学大学院工学系研究科の幾原雄一特別研究教授は「ARIM 事業の前身のナノテクノロジープラットフォーム事業では、ナノ計測とナノ加工を中心とした学内外および産業界の研究支援と共同研究が主な業務でした。2021年に発足した ARIM 事業では、ナノプラ事業のコンセプトを継承しつつこれより得られたデータを集積するとともにデータ構造化を行い、材料開発のための有用なツールを構築することを目的としています。本学では、mdx という極めて高性能の DX システムがすでに開発されており、ARIM 事業を mdx と一体化して運用することによって、今回の ARIM-mdx Data System を全国規模で推進することに目途が立ちました。今後は、本システムをさらに加速して運用し、我国のマテリアル革新力の強化に貢献していくことを期待しています。一方、物質・材料研究機構(NIMS)におけるデータ中核拠点とは連携してデータ共有を行なうことも計画しています。」とコメントしています。

発表者・研究者等情報

東京大学

情報基盤センター

華井 雅俊 特任助教

田浦 健次郎 教授

大学院工学系研究科

石川 亮 特任准教授

幾原 雄一 特別研究教授

論文情報

Masatoshi Hanai, Ryo Ishikawa, Mitsuaki Kawamura, Masato Ohnishi, Norio Takenaka, Kou Nakamura, Daiju Matsumura, Seiji Fujikawa, Hiroki Sakamoto, Yukinori Ochiai, Tetsuo Okane, Shin-Ichiro Kuroki, Atsuo Yamada, Toyotaro Suzumura, Junichiro Shiomi, Kenjiro Taura, Yoshio Mita, Naoya Shibata, Yuichi Ikuhara “ARIM-mdx Data System: Towards a Nationwide Data Platform for Materials Science” In Proceedings of 2024 IEEE International Conference on Big Data (IEEE BigData’24), To appear

関連情報

特許出願

特願 2023-156343 “IoT デバイス、データ転送システムおよびデータ転送方法”

ARIM-mdx Data System ウェブページ

<https://arim.mdx.jp/>

関連論文

Masatoshi Hanai, Mitsuaki Kawamura, Ryo Ishikawa, Toyotaro Suzumura, Kenjiro Taura, “Cloud Data Acquisition from Shared-Use Facilities in A University-Scale Laboratory Information Management System” In Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing 2023 (UCC’23), No 21, page 1–9

研究助成

本研究は以下の支援のもとで実施されました

- 文部科学省 マテリアル先端リサーチインフラ (ARIM Japan)
<https://nanonet.mext.go.jp/>
- データ活用社会創成プラットフォーム mdx
<https://mdx.jp/>
- 学際大規模情報基盤共同利用・共同研究拠点 (JHPCN)
<https://jhpcn-kyoten.itc.u-tokyo.ac.jp/ja/>
- 文部科学省 データ創出・活用型マテリアル研究開発プロジェクト 再生可能エネルギー最大導入に向けた電気化学材料研究拠点 (DxMT / DX-GEM)
<https://www.dx-gem.t.u-tokyo.ac.jp/wp/>
- 文部科学省・国立情報学研究所 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
https://www.nii.ac.jp/creded/nii_ac_jp_creded.html

用語解説

(注1) マテリアル先端リサーチインフラ (ARIM Japan) : 全国 25 の大学、研究機関が保有する最先端の計測、分析、加工プロセス設備とその技術・ノウハウを、産学官の全ての研究開発者に共用提供するプロジェクト

(注2) データ活用社会創成プラットフォーム mdx: 全国 11 の大学・研究機関からなるデータ活用社会創成プラットフォーム協働事業体が運営する、研究環境を用途に合わせてオンデマンドで短時間に構築・拡張・融合できる、データ収集・集積・解析のためのプラットフォーム

(注3) 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) : 国内の大学や研究機関の計算機システムやストレージを高速ネットワークで結んだ共用計算環境基盤

(注4) JHPCN (学際大規模情報基盤共同利用・共同研究拠点) : 北海道大学、東北大学、東京大学、東京科学大学、名古屋大学、京都大学、大阪大学、九州大学にそれぞれ附置するスーパーコンピュータを持つ8つの施設を構成拠点とし、東京大学情報基盤センターがその中核拠点として機能する「ネットワーク型」共同利用・共同研究拠点

問合せ先

(研究内容については発表者にお問合せください)

東京大学情報基盤センター データ科学研究部門
特任助教 華井 雅俊 (はない まさとし)

東京大学情報基盤センター広報 (担当: 大林)

東京大学大学院工学系研究科 広報室