

材料分野における実験・シミュレーションデータの 統合データプラットフォームの構築: 全国規模の運用にむけて

東大情報基盤センター 華井雅俊

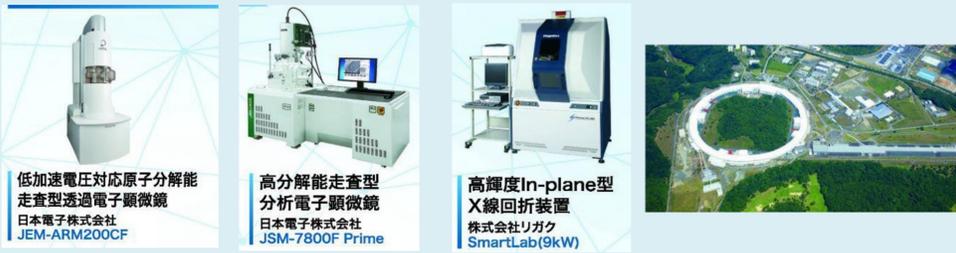


1. 背景: 文科省マテリアル先端リサーチインフラ (ARIM)

ARIM Japan

全国25大学・研究機関の材料実験設備共同利用プロジェクト

- 全国1000以上の共用装置
- アカデミックから企業まで幅広い利用
- 電子顕微鏡、X線回折装置、大型放射光、...
- 利用者からのデータ提供サービス



2. ARIM-mdx データシステム

- 全国の大型実験施設とスーパーコンピュータから創出される材料データ(実験データとシミュレーションデータ)をmdxの超大規模クラウドストレージに集約
- mdx上の高性能仮想化計算環境でデータ処理
- 各データサービス(NII GRDM, NIMS RDE)との連携

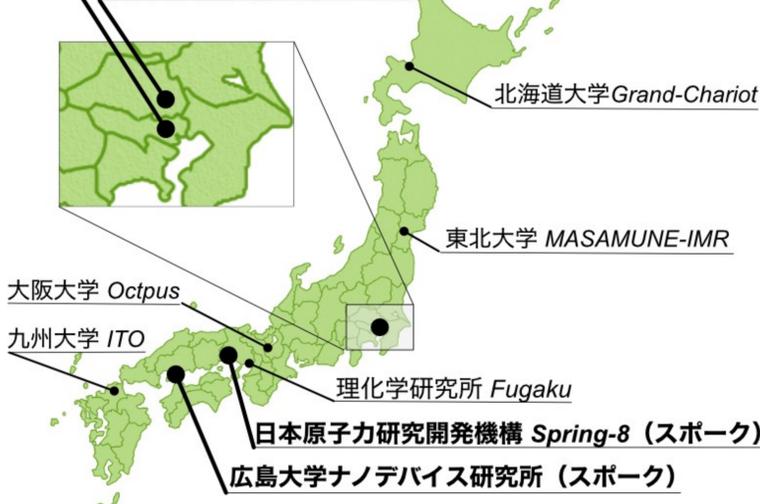
実験施設(東大、広大、JAEA)

スーパーコンピュータ群



東京大学マテリアル先端リサーチインフラ・データ拠点ハブ

東京大学情報基盤センター mdx / ARIM 超大容量ストレージ
東京大学情報基盤センター Wisteria
東京大学物性研究所 Ohtaka

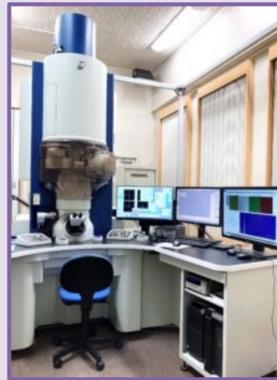


3. 非ネットワーク装置用データ転送IoTデバイス

- インターネットに繋がれない実験装置からの効率的なデータ転送
- IoTデバイスがUSBメモリをエミュレート
- ユーザー管理と連携し各自フォルダへ自動振り分け

装置制御用PC
(非ネットワーク)

独自IoTデバイス



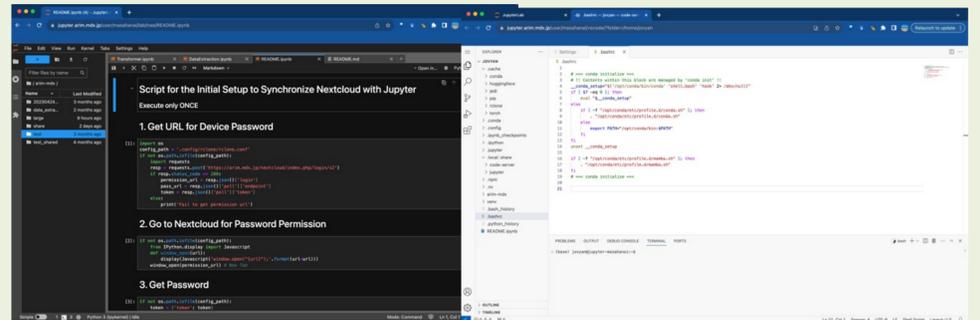
ARIM-mdx
データ
システムへ

USB

M. Hanai, M. Kawamura, R. Ishikawa, T. Suzumura, and K. Taura, The 16th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2023).

4. 高性能 Jupyterサービス

- mdx 上の仮想化計算環境をJupyterによって提供
- 転送したデータをmdx上で解析
 - Jupyter Notebookでお手軽データ解析
 - VScodeでガッツリ開発
 - Remote Desktop で GUI解析



5. 各種データサービスとの連携

- NII Gakunin RDM (研究管理サービス) とのデータフォルダ共有・連携
- NIMS RDE (材料データ公開システム) へのデータ提供



6. ユーザー数

- 今年度8月末から東大の装置利用者を中心に一般利用開始
 - **学外ユーザーや企業ユーザーを含めて現在753の利用ユーザー**
- 理論分野プロジェクトと実験分野プロジェクトが独立に利用中
 - **データ科学分野を融合した3分野の横断的な取り組みに注力**

