

(ポスター番号：P50)

(表題) Moonshot 目標 2 全体利用のための包括的未病データベースの構築

(所属機関) 国立情報学研究所

(氏名) 込山 悠介

(要旨文章)

Moonshot 目標 2 合原プロジェクトでは実験系研究と数理的研究が連携し未病データの解析・予測などの研究を実施している。特に「数理的連携研究、データベース構築および ELSI 支援体制構築」の課題においては、東京大学の藤原グループ(数理)と国立情報学研究所の山地グループ(データベース)の協働により、**図 1** の様な Moonshot 目標 2 全体で横断的に利用するための包括的未病データベースシステム(包括的未病 DB)の開発・構築を行っている。括的未病 DB は、**図 2** のように国立情報学研究所の研究データ基盤 NII Research Data Cloud (NII RDC)をベースに未病用に機能追加・拡張することで実現している¹。NII RDC は Moonshot プロジェクトにおける先端的な研究データ管理・利活用の指針に沿って機能開発・運用されており、標準的な研究データ管理の機能も具備している。包括的未病 DB は、非公開・制限共有で未病研究データを管理・共有するための基盤 GakuNin RDM²、それと連携して指定した計算機資源上で高度な計算を行うためのデータ解析機能 GakuNin Federated Computing Services (Jupyter)、そして、データをコミュニティで限定公開や、インターネット上で公開するための基盤 JAIRO Cloud で構成されている³。GakuNin RDM は研究データ管理 (RDM: research data management) サービスであり、利用機関の共同研究者とクローズドな領域でデータ管理・共有ができる大容量ストレージを備えたシステムである。GakuNin RDM には JupyterHub を用いたデータ解析機能があり、数理的研究におけるデータ分析環境とデータをセットで管理することができ研究再現性を高めている。JAIRO Cloud はリポジトリとよばれるタイプのシステムであり、データを公開するための業務プロセスを支援することに特化されている。キュレーターのためのファイルやメタデータの登録機能、利用者によるデータ検索とファイルプレビュー機能などを備える。不特定多数へのデータ公開用途以外でも、コミュニティ内での承認を伴う限定公開のための機能も持つ。包括的未病 DB は、ユーザインタフェースや DB 構造の設計に即しては、藤原グループで策定された未病データ用のメタデータを用いている。メタデータは実験系研究者からのフィードバックを受けている他、ヒトデータに関する記述については同じ課題に属する ELSI (Ethical, Legal and Social Issues) グループからも助言を得ている。今後の課題として、実際に各研究室での研究プロセスに沿った実務上の機能要件が十分かどうかを更に検証する必要がある。今後も DB のワーキンググループを通じ、実験系研究者へのヒアリングを継続的に実施し、システムの機能改善をおこなっていく。

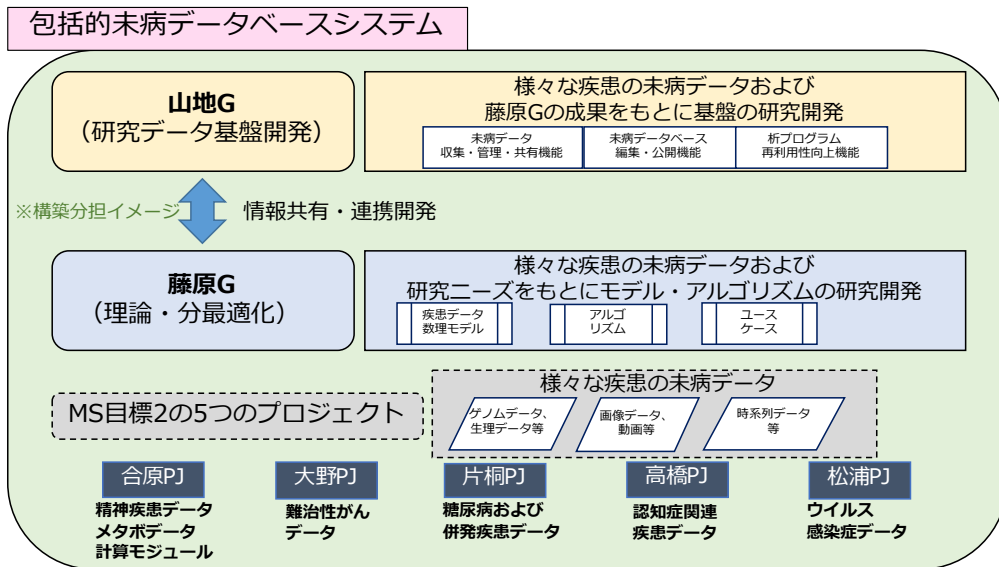


図1 包括的未病データベースシステムの構想

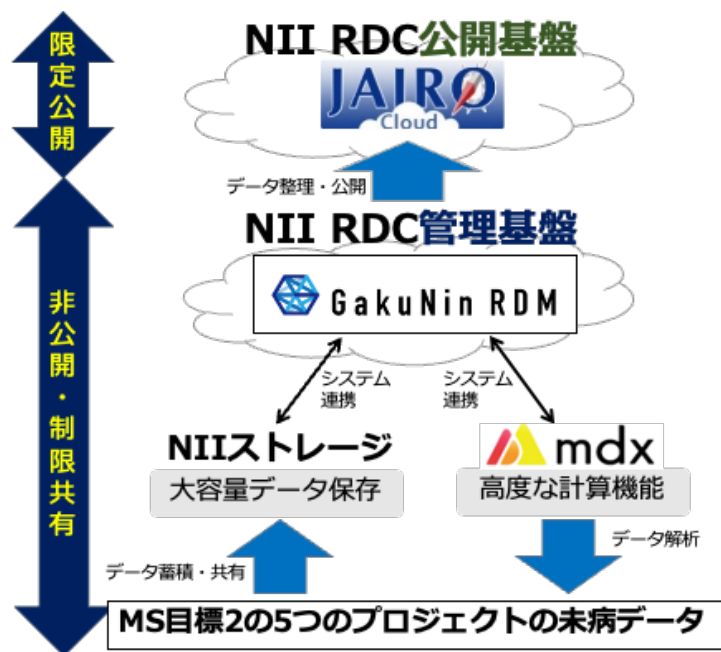


図2 包括的未病データベースシステムを構成する管理基盤と公開基盤

参考文献

1. Komiyama, Y., Hayashi, M., Kato, F., Ohmukai, I. & Yamaji, K. 学術機関に向けた研究データの管理と共有のための共通基盤の開発. in *IPSJ SIG-SPT 35 vols 2019-SPT-3 1-7* (2019).
2. Komiyama, Y. & Yamaji, K. Nationwide Research Data Management Service of Japan in the Open Science Era. in *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI) 129-133* (IEEE, 2017)..
3. 林正治, 林豊, 新妻聡 & 山地一禎. JAIRO Cloud とコミュニティ ーコミュニティ主導のクラウドサービスの実現ー. **2**, 32-46 (2021).