

実運用に基づくプロテオームデータ管理システムの メタデータ設計と外部リポジトリ連携



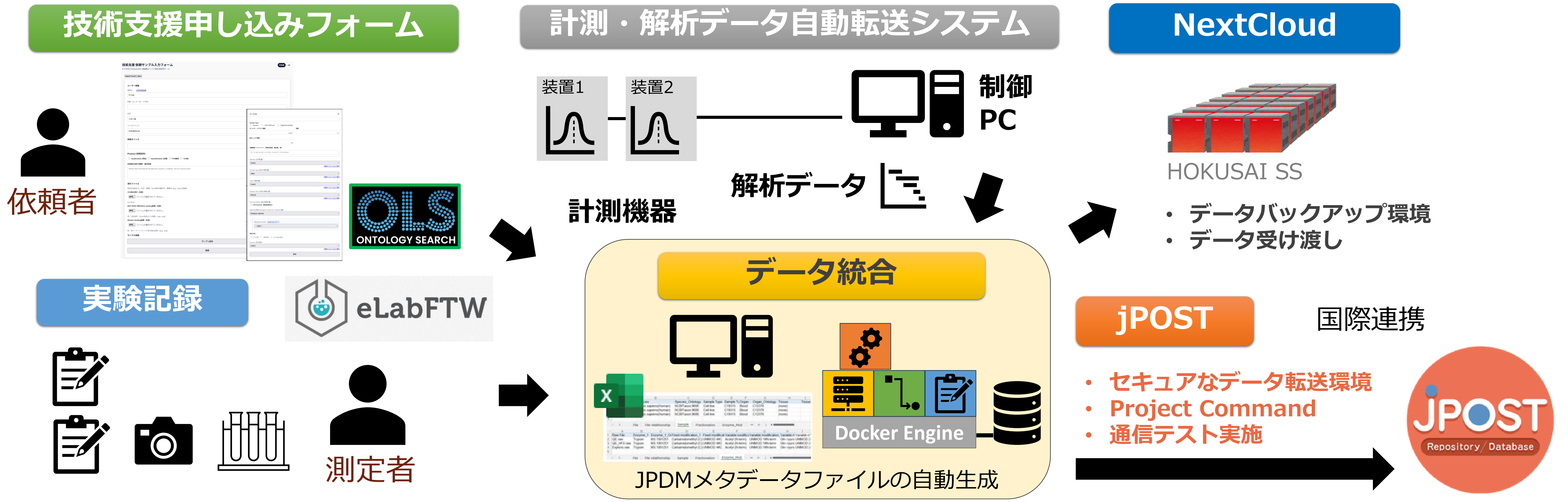
小池 仁美¹、東 裕介¹、前田 真秀¹、小松 輝久¹、鈴木 健裕²、大月 香³、中川 れい子⁴、堂前 直²、宮坂 信彦³、清水 義宏⁴、
俣賀 宣子¹、高橋 悠志⁵、吉沢 明康⁵、奥田 修二郎⁵、大浪 修一¹

¹理化学研究所 最先端研究プラットフォーム連携（TRIP）事業本部、²理化学研究所 環境資源科学研究センター、³理化学研究所 脳神経科学研究センター
⁴理化学研究所 生命機能科学研究センター、⁵新潟大学医学部 メディカルAIセンター

概要

プロテオームデータの再利用や統合解析には、測定データに加え、試料情報や解析条件を標準化されたメタデータとして管理することが重要である。本研究では、複数のファシリティで運用するデータ管理システムを対象に、実運用の知見をもとにメタデータ設計の改善を進めた。入力支援とメタデータの半自動取得により入力負担の軽減とデータ品質の向上を図るとともに、電子ラボノートとの連携、認証・アクセス制御およびバックアップ環境の整備を行い、複数拠点での運用基盤を構築した。本年度は、jPOSTチームが整備したデータ転送環境を用いた通信テストを実施し、連携に向けたメタデータ変換の基本方針の整理を進めた。本ポスターでは、これらの取り組みについて報告する。

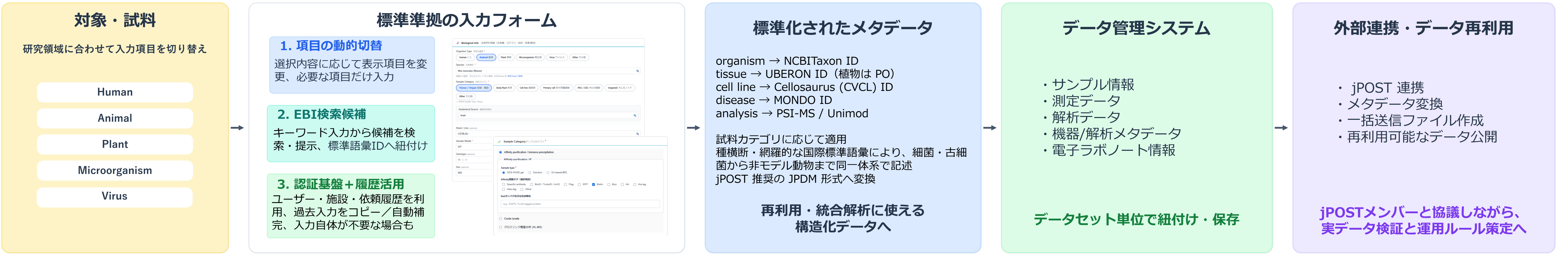
データ管理システム



試料の多様性に対応し、標準語彙IDへ自動で紐付ける入力フォーム

ユーザが持つ試料情報を入力するフォーム。試料の種類ごとに項目・語彙を切り替えられ、ファシリティごとに項目をカスタマイズできる。入力内容は標準語彙IDに紐付けて保持され、データ管理システム上で測定・解析データと対応付けることで JPDM 形式のメタデータが生成される。

研究者は使いやすいフォームで入力し、システム側で標準語彙・IDへ紐付ける



実運用に向けた基盤整備

実運用に必要な認証・データ保全の基盤を整備し、先行運用を開始した

認証基盤 — 所属を問わず利用できる
・理研内ユーザー：Shibboleth で直接認証
・外部ユーザー：学認経由（Orthros）で接続
理研内外のユーザーを共通の認証基盤で識別・管理

データ保全 — 失わない、待たせない
・バックアップ環境を構築し長期保存に対応
・依頼者へのデータ受け渡しを迅速化

ユーザー支援 — 二度目以降の入力を省く
・過去の依頼内容をコピーして再利用
・入力項目の自動補完



電子ラボノート elabFTW の活用

実験記録を電子化し、プロジェクト単位でデータと対応付けて管理

記録の対応付け — プロジェクトごとにラボノートを作成し、依頼・データセットと紐付けて管理

機能拡張 — 化学構造エディタ Ketcher を導入。構造式の描画・分子量計算に加え、ペプチドの構造作成・分子量計算にも対応

既存データの移行 — Excel で管理していた試料情報を一括で取り込み

継続的な更新 — オープンソースのため公式更新を取り込みつつ、必要な機能を独自に拡張できる

今後の展開

今後は、jPOSTメンバーと共同で実データを用いた検証と運用ルールの策定を進め、外部リポジトリへの登録機能の実装を目指す。あわせて、各ファシリティの要望を反映した入力フォームの改善、電子ラボノートからの試料情報の取り込み、LLM を活用した入力支援を進める。対応拠点を拡大し、プロテオームデータの再利用性向上に貢献していく。

