

2025 年度年次報告書

ライフサイエンスデータベース統合推進事業
統合化推進プログラム

【研究課題名】「バイオイメーシングデータのグローバルなデータ共有システムの構築」

研究代表者：

大浪 修一（理化学研究所 生命機能科学研究センター チームディレクター）

研究分担者：

なし



©2025 大浪 修一(理化学研究所) licensed under CC表示4.0国際

1. 研究進捗状況と成果

2025 年度は、「バイオイメーシングデータのグローバルなデータ共有システムの構築」に向けて、以下の取り組みを推進した。

データ登録の推進と利用拡大：SSBD:repository へのデータ登録呼びかけの対象を AMED-CREST 各領域や bioRxiv のプレプリントへと拡大した。また、第 63 回日本生物物理学会年会・第 98 回日本生化学会大会でのランチョンセミナーや関連シンポジウムの開催、Instagram を活用したアウトリーチ活動を実施した。その結果、2024 年度比で訪問者数約 30 倍・訪問数約 10 倍のアクセス増加を達成し、データ寄託件数も計 86 件に増加した。

グローバルなデータ共有基盤の整備：BioImage Archive (BIA)・Image Data Resource (IDR)・SSBD の三者間でのメタデータ統合に向けて、R0-Crate を採用したメタデータ記述方針を策定し、SSBD:repository および SSBD:database の合計 389 件のプロジェクトについて R0-Crate ファイルを作成した。また、REMBI 準拠の SSBD Ontology を整備し、SPARQL 横断検索および外部オントロジーとの連携を実証した。EMBL-EBI 主導のもと GIDE 相互検索ポータルのプロトタイプ開発も進展し、次年度初頭の正式リリースを見込んでいる。

データ解析・検索機能の強化：細胞追跡ツール ELEPHANT の GPU 実行環境を整備し、ハンズオンワークショップを開催した。また、OME-Zarr を基盤とした類似画像検索プロトタイプの開発と UMAP による表現型可視化システムの構築を進め、データ再利用性の向上を図った。

データベース運用基盤の整備：SSBD:database のメタデータモデルの改良とデータベースシステムの更新を行い、検索性能を飛躍的に向上させた。また、HOKUSAI SS において rustfs を用いた S3 互換ストレージを構築し、OME-Zarr データ提供の試行を通じてデータベースの理研共同利用計算システムへの移行に向けた基盤整備を推進した。

2. 代表的な成果事例

1. Onami, S., Kyoda, K., Yamagata, Y., Itoga, H. foundingGIDE Deliverable D6.1: Report in metadata model overlap and gaps. *Zenodo*, <https://doi.org/10.5281/zenodo.16794787> (2025).

概要：本論文の目的は、バイオイメーシングデータの相互運用性を向上させるために、各種データリソース間におけるメタデータ統合の要件を定義することである。これを達成するため、まず主要なバイオイメーシングデータリソースである BioImage Archive (BIA)、Image Data Resource (IDR)、SSBD:database のメタデータモデルを比較し、メタデータ構成要素の重複と相違点を明らかにするとともに、統合すべきメタデータ構成要素とオントロジーを特定した。次に、バイオイメーシングデータの再利用に関するさまざまなユースケースを分析し、データ再利用を促進するメタデータ統合の要件を明確にするため、必要なメタデータ構成要素、適切なオントロジー・統制語彙、および関連コミュニティを整理した。最後に、グローバルなバイオイメーシングデータ共有に向けたメタデータ統合の達成に向け、BIA・IDR・SSBD それぞれのマイルストーンを提示した。さらに、より持続可能で柔軟なバイオイメーシングデータ統合の枠組みを検討するため、メタデータの管理・共有手法についても概説した。

以上