

2023年度 研究開発実施報告

概要

研究開発課題名	バイオイメーキングデータのグローバルなデータ共有システムの構築
開発対象データベースの名称(URL)	SSBD データベース (https://ssbd.riken.jp)
研究代表者氏名	大浪 修一 (50348843)
所属・役職	理化学研究所 生命機能科学研究センター チームリーダー (2024年3月時点)



目次

概要1

目次2

§1. 研究実施体制	4
§2. 研究開発対象とするデータベース・ツール等	4
(1) データベース一覧	4
(2) ツール等一覧	4
§3. 実施内容	5
(1) 本年度の研究開発計画と達成目標	5
1a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上 5	
1b) SSBD:repository の利用者の拡大	5
1c) SSBD:database のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上	6
2a) リポジトリのための最小限のメタデータの標準化や API の共同開発等を含む、SSBD:repository と BioImage Archive とのデータ共有の実現	6
2b) 高付加価値データベースのための次世代のファイルフォーマットやメタデータの管理システム、API の 共同開発等を含む、SSBD:database と IDR とのデータ共有の実現	7
2c) グローバルなデータ共有のためのファイルフォーマット、メタデータ、API の標準化等を含む、グロー バルなデータ共有システムの構築	8
3a) バイオイメージングデータに関する新たな情報解析手法の開発のためのデータ基盤の整備	8
3b) バイオイメージングデータに関する主要な画像情報処理ソフトウェアの収集と実行環境の提供	9
(2) 進捗状況	10
1a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上 10	
1b) SSBD:repository の利用者の拡大	10
1c) SSBD:database のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上 11	
2a) リポジトリのための最小限のメタデータの標準化や API の共同開発等を含む、SSBD:repository と BioImage Archive とのデータ共有の実現	11
2b) 高付加価値データベースのための次世代のファイルフォーマットやメタデータの管理システム、API の 共同開発等を含む、SSBD:database と IDR とのデータ共有の実現	12
2c) グローバルなデータ共有のためのファイルフォーマット、メタデータ、API の標準化等を含む、グロー バルなデータ共有システムの構築	12
3a) バイオイメージングデータに関する新たな情報解析手法の開発のためのデータ基盤の整備	13
3b) バイオイメージングデータに関する主要な画像情報処理ソフトウェアの収集と実行環境の提供	13
4a) SSBD データベースの運用の理化学研究所共同利用計算システムへの移行	13
§4. 成果発表等	14
(1) 原著論文発表	14
① 論文数概要	14
② 論文詳細情報	14
(2) その他の著作物(総説、書籍など)	15
(3) 国際学会および国内学会発表	15
① 概要	15
② 招待講演	15
③ 口頭講演	16
④ ポスター発表	16
(4) 知的財産権の出願 (国内の出願件数のみ公開)	17
出願件数	17
(5) 受賞・報道等	17
① 受賞	17

②	メディア報道.....	17
③	その他の成果発表.....	18
§5.	主要なデータベースの利活用状況.....	19
(1)	アクセス数.....	19
①	実績.....	19
②	分析.....	19
(2)	データベースの利用状況を示すアクセス数以外の指標.....	19
(3)	データベースの利活用により得られた研究成果(生命科学研究への波及効果).....	20
(4)	データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果).....	20
§6.	研究開発期間中に主催した活動(ワークショップ等).....	21
(1)	進捗ミーティング.....	21
(2)	主催したワークショップ、シンポジウム、アウトリーチ活動等.....	23

§1. 研究実施体制

グループ名	研究代表者または主たる共同研究者氏名	所属機関・役職名	研究題目
大浪グループ	大浪 修一	理化学研究所・生命機能科学研究センター	バイオイメージングデータのグローバルなデータ共有システムの構築

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	SSBD データベース	SSBD:database	https://ssbd.riken.jp

【その他のデータベース】

No.	名称	別称・略称	URL
1	SSBD リポジトリ	SSBD:repository	https://ssbd.riken.jp/repository/

(2) ツール等一覧

No.	名称	別称・略称	URL
1			

§3. 実施内容

(1) 本年度の研究開発計画と達成目標

1a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:repository は、最小限に絞った必須メタデータとともにデータを公開するサービスである。最小限であっても多岐にわたるメタデータを、実験を実施するデータの産出者や論文の著者から適切に獲得するために、これまで、広く一般に利用されている表計算フォーマット(Microsoft Excel 形式)のテンプレートを用いてきた。この仕組みには、誤りを入力時に発見することが難しい、何らかの手段でファイルのやり取りが必要になる、ファイルからデータベースへのデータ登録作業が必要になるなどの問題がある。そこで、適切なウェブフォームを用いることで、ファイルのやり取りをなくし、誤りによる作業の手戻りやデータ登録の作業を軽減する。人手を介さないことからスケーラビリティが向上する。さらに、頻出入力例や利用者ごとの入力履歴から、入力内容を自動的に補完する機能を追加し、ユーザーエクスペリエンスを向上させる。

第 1 年度は、1) SSBD:repository のデータ登録システムに利用者および管理スタッフが利用可能なウェブフォームによるデータ入力システムを試作した。2) メタデータの入力不足、余分、誤りを発見するスクリプトを実装し、表計算フォーマットのテンプレートに対して適用し有効性を確認した。

第 2 年度は、データ入力システムを運用し、利用上の問題などを明らかにする。また、メタデータの入力不足、余分、誤りを発見するスクリプトを入力システムに組み込み、ウェブフォームでの有効性を確認する。また、試作したデータ入力システムに頻出入力例から入力を補完する仕組みを実装し、使用者を補助する仕組みを構築する。

1b) SSBD:repository の利用者の拡大

新学術領域「先端バイオイメージング支援プラットフォーム(ABiS)」などのイメージングを支援するプロジェクト、新学術領域「シンギュラリティ生物学」で開発されている AMATERAS など最先端のイメージング装置を開発し利活用を推進している施設、JST-CREST、AMED-CREST などと連携し、そこで得られた画像データを SSBD:repository に登録してもらうことにより、利用者の拡大をはかる。また、国立研究開発法人日本医療研究開発機構(AMED)と連携し、AMED が支援する研究から得られたデータのうち、イメージング関連データを SSBD:repository で引き受けることにより、データを効率的に格納・共有するスキームを確立する。

SSBD:repository はオリジナルデータを迅速に公開するサービスであることから、顕微鏡画像や生命動態定量データをオリジナルファイルのまま公開しダウンロードさせているが、顕微鏡画像管理システム OMERO の画像自動登録機能による画像プレビューや、一定の基準の生命動態定量データについて簡単な可視化を行うことで、サービスの利便性を向上させ利用者の拡大を図る。

また、SSBD:repository ではデータセットが公開されるごとに SNS (twitter) でアナウンスを行っており、論文やプレプリントの著者によりリツイートされることで論文とデータセットだけでなく SSBD:repository の宣伝にもなっている。これを Image.sc などの専門的な SNS にも展開することで利用者の拡大を図る。

第 1 年度は 1) JST-CREST「多細胞」領域で取得された画像、新学術領域「シンギュラリティ生物学」で開発されている AMATERAS による画像などの SSBD:repository への登録を推奨する取り組みを実施した。また、先端バイオイメージング支援プラットフォーム ABiS の支援を受けて計測されたデータの SSBD:repository への共有を誘導する仕組みを構築した。2) SSBD:repository に顕微鏡画像管理システム

OMERO を導入し、顕微鏡画像のプレビューを実現した。

第 2 年度は、ABiS により支援されているプロジェクトの画像の登録を推奨する取り組みを実施するとともに、AMED に関連した医科学系の画像の登録について検討する。さらに、生命科学及び医科学に関する主要な学会と連携した「SSBD アドバイザリーボード(仮称)」を設立し、各学会からの画像の登録を進めるとともに、SSBD:repository から SSBD:database に選出されるデータの品質の保証についての意見を聴取し、品質と再利用性に関するガイドラインを作成する。

1c) SSBD:database のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:database は十分なメタデータとともにデータを公開するサービスである。メタデータは内容が多岐にわたり、また数も多いため、キュレーターによるデータの入力・データ産出者や論文の著者による内容の確認と修正・データベースへのデータの登録のために、これまで、広く一般に利用されている表計算フォーマット (Microsoft Excel 形式) のテンプレートを用いてきた。この仕組みには、オントロジー等の正確な表現が必要なデータの誤りを見逃す、データの確認が煩雑かつデータベース登録後のメタデータの扱いや表示に結びつかない、何らかの手段でファイルのやり取りが必要になる、ファイルからデータベースへの登録作業が必要になるなどの問題がある。そこで、適切なウェブフォームと確認用のウェブサイトを用いることで、ファイルのやり取りをなくし、メタデータの誤った登録を防止する。キュレーターは多数のメタデータを入力することから、画面遷移を最小とし、同じデータの繰り返し入力を削減するユーザインタフェースとする。ファイルのやり取りを削減することと合わせて、作業のスケーラビリティを拡大させる。さらに、認証付きのメタデータ確認用ウェブサイトを用意し、データ産出者や論文の著者にデータベース登録後のイメージを伝えるとともに、連絡先などの非公開情報を安全に確認する手段とすることでユーザーエクスペリエンスを向上させる。

第 1 年度は、1) SSBD:database のデータ登録システムに管理スタッフが利用可能なウェブフォームによるデータ入力システムを試作した。2) メタデータの入力不足、余分、誤りを発見するスクリプトを実装し、表計算フォーマットのテンプレートに対して適用し有効性を確認した。さらにキュレーターによるデータ登録のスケーラビリティを向上させるために、登録候補となる論文情報を収集するデータベースシステムを SSBD の一部として新たに構築した。生命科学・医科学の論文データベースである NCBI PubMed および PubMed Central データベースを、SSBD に関連するプロジェクトの助成番号、SSBD に関連するプロジェクトの研究者名と所属組織名で検索し、整理のためのタグ情報を付与して蓄積する。蓄積された論文情報は SSBD から直接参照可能であり、キュレーターによる収集作業を省力化するとともに、従来の数十倍の論文を整理して扱うことができるようになった。

第 2 年度は、データ入力システムを運用し、利用上の問題などを明らかにする。また、メタデータの入力不足、余分、誤りを発見するスクリプトを入力システムに組み込み、ウェブフォームでの有効性を確認する。また、試作したデータ入力システムに頻出入力例から入力を補完する仕組みを実装し、使用者を補助する仕組みを構築する。また、新たに作成した論文収集データベースを改良し、researchmap などを用いて日本から発信された論文をより効率的に収集する仕組みとする。

2a) リポジトリのための最小限のメタデータの標準化や API の共同開発等を含む、SSBD:repository と BioImage Archive とのデータ共有の実現

BioImage Archive は、Global BioImaging が提唱するバイオイメーキングデータのグローバルなデータ共有システムにおけるデータリポジトリの機能を提供するために、2019 年に EMBL-EBI に設立されたバイオ

イメージングデータのデータリポジトリである。SSBD:repository と BioImage Archive の連携については、BioImage Archive の開発チームの代表である EMBL-EBI の Matthew Hartley 博士が 2020 年初頭に EMBL-EBI に着任して以降、Harley 博士と大浪の間で継続的に議論が行われている。両データリポジトリの間でデータを共有するためには、各リポジトリにおいてデータに付与するメタデータの標準化が必須である。BioImage Archive と SSBD データベースは 2019 年よりメタデータの標準化を目指した共同研究を開始しており、2019 年からの議論をまとめた、現時点での推奨されるメタデータ (REMBI) に関する論文を 2021 年に発表した (Sarkans et al., Nature Methods 18, 1418-1422, 2021)。本研究開発では、これまでの議論を土台にして両データリポジトリ間のメタデータ標準化の議論を加速し、両データリポジトリが採用する標準メタデータを確立する。さらに、メタデータのデータリポジトリへの登録を支援するユーザー支援システムを共同開発し、メタデータを利用して両データリポジトリより共有されているバイオイメージングデータの横断検索が可能な、統合データ検索システムを構築する。

第 1 年度は、1) SSBD:repository のメタデータを、Sarkans et al., 2021 で推奨された REMBI 規格に準拠させる改良を開始した。メタデータの変更にはメタデータの再入力や再キューションが必要となるため、新たなメタデータでの運用は順次開始する。2) SSBD:repository のメタデータと BioImage Archive のメタデータの連携について、BioImage Archive を主宰する Matthew Hartley 氏および BioImage Archive の基幹システムとして利用されている OMERO を開発する Open Microscopy Environment (OME) チームとオンラインおよび現地での議論を開始した。

第 2 年度は、メタデータの連携や共通化について、新たに Euro-Bioimaging プロジェクトの関係者を加え、具体的なメタデータの内容や実装方法について議論するとともに、SSBD システムと BioImage Archive との共通の検索機構を試作する。

2b) 高付加価値データベースのための次世代のファイルフォーマットやメタデータの管理システム、API の共同開発等を含む、SSBD:database と IDR とのデータ共有の実現

Image Data Resource (IDR) は、バイオイメージングデータの管理のためのソフトウェアプラットフォームとして世界のデファクト標準となっている OMERO を開発する OME コンソーシアムが EMBL-EBI と共同で 2017 年に設立したバイオイメージングデータの高付加価値データベースである。SSBD データベースは 2013 年の設立以来、バイオイメージングデータの管理には OMERO を採用しており、OME コンソーシアムとは長期間にわたり密接な共同研究を続けている。IDR および OME コンソーシアムの代表の University of Dundee の Jason Swedlow 教授と大浪は、10 年以上共同研究を続けており、IDR と SSBD:database の連携についても継続的な議論を行ってきた。実験生物学者から情報科学者/AI 研究者までの幅広い研究者による利活用が期待される高付加価値データベースには、デスクトップ PC を用いた可視化解析からクラウド環境での大規模データ解析までの様々な計算環境において高いパフォーマンスを示すファイルフォーマットによるデータの管理が求められる。また、分野ごとに利活用求められるメタデータの項目が大きく異なる状況下で、多様に付与された豊富なメタデータの統合を可能にする柔軟なメタデータの管理システムが求められる。IDR と SSBD データベースは、このような次世代のデータ管理システムの開発を目指した共同研究を開始しており、これまでの議論をまとめた次世代型のファイルフォーマットとメタデータ管理システムに関する論文を 2019 年に発表した (Moore et al., SWAT4HCLS 2019, 44-53, 2019)。本研究開発では、これまでの議論を土台にして、次世代のファイルフォーマットとメタデータ管理システムの開発を加速し、豊富なメタデータを利用して両データベースより共有されているバイオイメージングデータの高レベルな横断検索が可能であり、幅広い計算環境における多様な解析に最適化されたフォーマットでデータが提供される、SSBD:database と IDR の統合データ検索システムを構築する。

第1年度は、2023年2月28日-3月4日の日程でSSBDチーム5名がUniversity of Dundee(英国・ダンディー市)を訪問し、Jason Swedlow 教授が率いる IDR/OMERO 開発チームと SSBD:database と IDR との協働に関する打ち合わせを行い、今後のメタデータの連携や共通化を目指した開発計画について合意した。

第2年度は、メタデータの連携や共通化について、1) 2023年2月のSSBDチームによる英国・ダンディー市の訪問において具体化された IDR および OMERO チームとの協働について、1)メタデータの共通化および付与されていないメタデータの付与し、2)さらにオントロジーを用いて付与されていないメタデータを補完するハーモナイズ法について具体化する。

2c) グローバルなデータ共有のためのファイルフォーマット、メタデータ、API の標準化等を含む、グローバルなデータ共有システムの構築

SSBD で使用するデータ共有のためのファイルフォーマットは、バイオイメージング分野の国際的なコミュニティで合意したフォーマットに従うものとする。画像データのフォーマットは、欧州の OME チームを中心に開発が進められている OME-NGFF を採用することとし、共通のメタデータを持つ TIFF, HDF5, Zarr を基盤としたフォーマットで共有する方針である。定量データのフォーマットについては、これまでに BDML, BD5 といったフォーマットを開発してきたが、これらを開発したノウハウを活かして、画像データの次世代フォーマットである OME-NGFF と親和性の高いフォーマットを、国際的なコミュニティの同意を得ながら開発する予定である。これらのフォーマットにより格納されたデータに対する API も、フォーマットと同様に、国際的なコミュニティと共同で開発する予定である。共通のファイルフォーマットをグローバルに採用することにより、API の開発やツール群の開発の効率化が期待される。さらに、SSBD では、国際的なメタデータ標準化を目指したコミュニティ、QUAREP LiMi により登録が推奨されているメタデータを採用する。

第1年度は、定量データの次世代フォーマットの初版である BD-zarr を開発し、国際的なバイオイメージングコミュニティのフォーラムである image.sc にて提案を行った。いくつかのフィードバックを元に、現在フォーマットの改良版を作成している。また、SSBD:database で共有している時空間情報を含むデータに対する BD5 ファイルを可視化するデスクトップアプリケーションを開発した。

第2年度は、引き続き BD-Zarr の開発を行い、次世代のバイオイメージングファイルフォーマットである OME-Zarr の表形式の新しい仕様に対応させた改良版を開発する。本フォーマットのコミュニティへの認知度を上げるために、国際的なバイオイメージングフォーマット開発の中心である OME-NGFF コミュニティのミーティングやハッカソン等の集まりに参加し、発表、提案を積極的に行うことにより、次世代バイオイメージングファイルフォーマットへの貢献をはかる。また、試用を含めたフォーマット利用を促進するために、napari や ImageJ など既存ツールのいずれかに対応させる。また、SSBD:database のメタデータについては、バイオイメージングコミュニティが推奨するメタデータである REMBI や顕微鏡の構成や設定に対する標準化を目指している QUAREP LiMi が推奨するメタデータに対応させる。さらに、欧州のバイオイメージングデータの高付加価値データベースである IDR との連携を開始し、遺伝子情報や GO (Gene Ontology)を利用した相互検索、解析を可能にする基盤を構築する。

3a) バイオイメージングデータに関する新たな情報解析手法の開発のためのデータ基盤の整備

深層学習を含む情報処理解析技術の開発を促進するために、多次元データへの並列アクセスを可能にする次世代のデータ形式で、画像データおよび定量データを提供する。具体的にはクラウドからネイティブにア

アクセス可能で、柔軟で包括的なメタデータ構造を持つ Zarr を基盤としたデータ形式で、データを公開することにより、解析ツールからのデータアクセスを高速かつ効率的に行えるようにする。TIFF や HDF5 など確立されたデータ形式で格納されたデータファイルとの相互変換ツールを準備することにより、迅速かつ持続可能なデータ提供をはかる。画像データの次世代データ形式である OME-NGFF においては、既に bioformats2raw, raw2ometiff が準備されており、これらのツールを利用する。定量データについても同様の方針で、変換ツールを準備する予定である。

第 1 年度は、SSBD:database で共有されている画像データの OME-Zarr 変換に対するフィージビリティスタディを行った。既存ツールである bioformat2raw を利用することにより、OME-Zarr への変換が可能であることを確認した。また、メタデータについては、ファイル自体にそれを含むデータに対しては、自動でメタデータを抽出して OME-Zarr のメタデータとして出力できることがわかった。しかしながら、SSBD:database には、メタデータを抽出できないファイルが相当数、存在することが判明した。これらのデータに対しては、SSBD:database に含まれるメタデータを利用して、OME-Zarr のメタデータファイルを作成する必要がある。

第 2 年度は、第 1 年度の結果を踏まえて、SSBD:database で共有されているデータファイルの OME-Zarr への変換を行う。また、メタデータの確認を行い、必要に応じて OME-Zarr ファイルセットに付属するメタデータの編集を行う。さらに、外部公開に向けた取り組みとして、理化学研究所の共同利用計算システムで運用されているオブジェクトストレージの利用を検討する。

3b) バイオイメージングデータに関する主要な画像情報処理ソフトウェアの収集と実行環境の提供

SSBD では、これまでに画像データと定量データを収集・共有を行ってきた。NBDC 主催の分子生物学会でのブースやバイオイメージングコミュニティにおけるユーザーからの要望の多くは、画像データから定量データを計測する要望であった。

この要望に応えるために、本開発期間には、新たに論文で利用された画像解析プログラムを Notebook 形式で収集・共有する。画像データの種類や目的に応じて、利用される画像処理ソフトウェアは異なる場合があるため、バイオイメージング分野における主要なソフトウェア(例えば、Fiji, CellPose, ilastik など)を利用したプログラムを揃える予定である。これにより、画像データ解析結果を再現したり、改変・改良を行ったりすることにより、新たな画像データへの適用が可能になる。

各種プログラムがどのような画像データを利用し、どのような解析結果が得られるのかを容易に理解できるようにするため、Notebook 形式の実行環境の整備も行う。これは、実行環境を整えることが困難な実験生物学者の要望にも応えることにもつながる。最初に、Google Colab や Binder といった無料で利用できる外部の実行環境を使ったサービスを提供する。これらの環境は、比較的軽い処理に適しており、サービス利用者、提供者双方に負担なく、気軽にプログラムを試すことが可能になる。しかしながら、論文で使用される画像処理には比較的重い処理も含まれることが予想される。このようなケースにも対応するために、独自にサーバを準備して BinderHub などを用いることにより、画像処理等を行うことができる実行環境を提供する。

また、SSBD で共有されている画像データは、深層学習用のトレーニングデータとしての利用が期待される。しかしながら、SSBD で共有されている画像データは最先端の技術で撮影されたものが多く、深層学習に利用するにはデータ数が充分でない場合も予想される。そこで、既存のデータ拡張の手法を用いることにより、データ数を増やすことでトレーニングデータとしての活用を促進する。

第 1 年度は、論文で開発または使用された画像処理ソフトウェア、ワークフローについての情報を収集した。

国内外で利用されている画像処理ソフトウェアとしては、Fiji、CellPose、CellProfilerなどが利用されており、また、独自開発された画像処理ワークフロー等については、GitHubなどのリポジトリでコードが公開されていた。これらのソフトウェアおよびワークフローは、Python上で動作させることが可能であり、BinderHub等を利用したJupyter Notebookの実行環境を整備することが望ましい。

第2年度は、これらを踏まえて、BinderHub等を利用した実行環境を構築し、これまでに論文で開発または使用された画像処理ソフトウェアやワークフローの動作を確認する。また、公開されているGitHub上のワークフローを、Google ColabやBinderといった外部の実行環境で動作させるサービスの試験運用を開始する。さらに、ワークフローを実行する環境として、国立情報学研究所(NII)が開発しているRDC (Research Data Cloud)で提供されているデータ解析機能(<https://rcos.nii.ac.jp/service/cs/>)の利用を検討する。

(2) 進捗状況

1a) SSBD:repository のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:repository のデータ登録システムに、利用者および管理スタッフが利用可能なウェブフォームによるデータ入力システムの運用を管理スタッフの中で試行した。特に、プレプリントや論文の情報の入力においては、1c) SSBD:database において先行して開発を進めた、登録候補となる論文情報を収集するデータベースシステムと連携し、プレプリントの公開サービスである bioRxiv などの文献情報の入力と管理を支援する仕組みを運用し、これらのシステムやデータの誤りを発見するスクリプトの問題を検出し、さまざまな入力に対応できるように改良を行った。

本年度は更にこれらのシステムを順次公開し、実際の利用者からのフィードバックを受ける計画であったが、後述するように、foundingGIDE プロジェクトの開始に伴い、「2a) SSBD:repository と BioImage Archive との連携」、および「2b) SSBD:database と IDR との連携」において、メタデータの統合やハーモナイズのためにメタデータの項目を大きく変更したため、これらに対応するメタデータの入力について見直す必要が生じた。そのため、表計算フォーマットを更新し利用者からのフィードバックや入力例を獲得することを優先し、ウェブフォームの公開は変更された項目についての情報が得られたのちに行うこととした。

1b) SSBD:repository の利用者の拡大

JST-CREST「多細胞」領域で取得された画像、新学術領域「シンギュラリティ生物学」で開発されている AMATERAS による画像などの SSBD:repository への登録を推奨する取り組みを実施した。2023 年度は JST-CREST「多細胞」領域に関連する著者を含む論文のデータが3件、新学術領域「シンギュラリティ生物学」に関連する著者を含む論文のデータが3件登録された。

ABiS により支援されているプロジェクトの画像の登録を推奨する取り組みにおいては、同プラットフォームでデータ共有に利用可能なデータプラットフォームを整備するなどの取り組みを行った。また、ABiS-GBI 2023 course – Image data: image analysis, data management and reuse (2023 年 7 月 3-7 日:岡崎、神戸)を ABiS および GBI と共同で開催し、ABiS の支援で取得されたデータの SSBD:repository への登録について広く宣伝を行った。

AMED に関連した医科学系の画像データについては、後述の SSBD アドバイザリーボードの設立に向けた議論の中で、日本癌学会、日本病理学会との協力関係が構築され、今後の具体的な連携方法についての

議論を開始した。また、後述の **foundingGIDE** プロジェクトにおいて、前臨床イメージングデータ・リポジトリの日欧豪間での統合が目標の一つとされたことから、我が国の前臨床イメージングに関するデータベース関係者との議論を開始した。

SSBD アドバイザリーボードの設立については、国内の生命科学系および医科学系のイメージングに関連する主要な学会11学会と **ABiS** に対し主旨説明等を行い、各学会等との協力関係の構築に成功した。各学会等からアドバイザリーボードメンバーの推薦を受け、第 1 回アドバイザリーボード会議の開催に向けての準備を開始した。

1c) **SSBD:database** のデータ登録システムのスケーラビリティの拡大とユーザーエクスペリエンスの向上

SSBD:database のデータ登録システムに、利用者及び管理スタッフが利用可能なウェブフォームによるデータ入力システムの運用を試行した。特にキュレーターによる論文の選択や入力においては、登録候補となる論文情報を収集するシステムを **SSBD** の一部として運用するとともに、これらのシステムやデータの誤りを発見するスクリプトの問題を検出し、さまざまな入力に対応できるように、改良を行った。

本年度はこれらのシステムを順次公開し、実際の利用者からのフィードバックを受ける計画であったが、後述するように、**foundingGIDE** プロジェクトの開始に伴い、「2a) **SSBD:repository** と **BioImage Archive** との連携」、および「2b) **SSBD:database** と **Image Data Resource** との連携」において、メタデータの統合やハーモナイズのためにメタデータの項目を大きく変更したため、これらに対応するメタデータの入力について見直す必要が生じた。そのため、表計算フォーマットを更新し利用者からのフィードバックや入力例を獲得することを優先し、ウェブフォームの公開は変更された項目についての情報が得られたのちに行うこととした。

2a) リポジトリのための最小限のメタデータの標準化や API の共同開発等を含む、**SSBD:repository** と **BioImage Archive** とのデータ共有の実現

SSBD:repository のメタデータを Sarkans et al., 2021 で推奨された **REMBI** 規格に準拠させる改良を行い、**SSBD metadata v3** として定義した。具体的な手順としては、**REMBI** で提案されたメタデータの枠組みと、**SSBD:repository** における必須メタデータおよび **SSBD:database** におけるメタデータを比較し、不足しているメタデータについて導入が可能かを検討し、観察対象の遺伝子情報、処置の情報、顕微鏡で撮影した際の対象の情報などを導入し、それらに適切なオントロジーを付与することとした。

2023年7月に **BioImage Archive** の主宰者である **Matthew Hartle** 博士およびチームメンバーである **Ayubuke Yoldas** 博士を、理化学研究所生命機能科学センター(兵庫県神戸市)に招き、**SSBD** との連携およびデータ共有の実現について打ち合わせを行った。**BioImage Archive** で使用されているメタデータ登録システム **BioStudies** の詳細について情報交換するとともに、最小メタデータの標準化や共同での検索システムの開発について合意した。

欧州の **Horizon Europe** 助成事業において、バイオイメージングのリポジトリの日欧豪の連携に関するプロジェクト **foundingGIDE** が採択され、2024年1月よりプロジェクトが開始された。当プロジェクトでは、バイオイメージングのデータリポジトリとして **SSBD**(日本)、**BioImage Archive** および **IDR**(欧州)のメタデータレベルでの連携を目標とし、更に前臨床画像データの日欧豪間の連携を目標としている。このため、「2a)リポジトリのための最小限のメタデータの標準化や API の共同開発等を含む **SSBD:repository** と **BioImage Archive** とのデータ共有の実現」は、**foundingGIDE** プロジェクトの取り組みを取り入れ、日欧豪の協働によって実施

することとなった。

2b) 高付加価値データベースのための次世代のファイルフォーマットやメタデータの管理システム、API の共同開発等を含む、SSBD:database と IDR とのデータ共有の実現

SSBD:database のメタデータを、IDR と共通化するための改良を行い、1)既存のメタデータの記述から遺伝子情報およびタンパク質情報を機械的および手作業を組み合わせで抽出し、112 プロジェクトにのべ 254 個の遺伝子名と 207 個のタンパク質名を付与した。新たなメタデータとして付与した。この改良により IDR とのメタデータの互換が可能となったが、既存のメタデータ記述が自由文であったことからメタデータの正確性と完備性が不足していることが判明した。そこで 2)IDR との連携も視野に入れた SSBD metadata v3 を定義した。この SSBD metadata v3 を用いて、生命科学とオントロジーの知識を十分に備えたキュレーターにより、既存のデータセットについて再アノテーションを行った。2024 年 3 月末までに 55 プロジェクト・338 データセットに対してのべ 430 個の遺伝子とタンパク質に関する情報、233 個の薬剤投与等の外部刺激に関する情報が新たに付与された。

2023 年 7 月に IDR および OME のチームメンバーである Jean-Marie Burel 博士および Petr Walczysko 博士を、理化学研究所生命機能科学研究センター(兵庫県神戸市)に招き、SSBD との連携およびデータ共有の実現について打ち合わせを行った。IDR 側より GO を用いたメタデータのハーモナイズ法について、疾患情報につながる新たな提案があり、メタデータの共通化や協働での検索システムの開発について合意した。

欧州の Horizon Europe 助成事業において、バイオイメージングのリポジトリの日欧豪の連携に関するプロジェクト foundingGIDE が採択され、2024 年 1 月よりプロジェクトが開始された。当プロジェクトでは、バイオイメージングのデータリポジトリとして SSBD(日本)、BioImage Archive および IDR(欧州)のメタデータレベルでの連携を目標とし、更に前臨床画像データの日欧豪間の連携を目標としている。このため、「2b) 高付加価値データベースのための次世代のファイルフォーマットやメタデータの管理システム、API の共同開発等を含む、SSBD:database と IDR とのデータ共有の実現」は、foundingGIDE プロジェクトの取り組みを取り入れ、日欧豪の協働によって実施することとなった。

2c) グローバルなデータ共有のためのファイルフォーマット、メタデータ、API の標準化等を含む、グローバルなデータ共有システムの構築

当初の計画を前倒し、画像データ解析により計測された定量データの次世代フォーマット、BD-Zarr の開発を行った。バイオイメージングデータの次世代フォーマットの OME-Zarr で新たに策定された表形式の仕様に準拠するよう BD-Zarr の改良を行い、BD-Zarr の初期バージョンの開発を完了させ、GitHub 上で公開した(<https://github.com/opensbd/bdz>)。また、Python ベースのバイオイメージングデータ可視化・解析ツールである napari 上で、OME-Zarr と BD-Zarr で格納したデータを可視化するツールのベータ版を開発した。OME-NGFF コミュニティが開催する Community Call に参加し、OME-Zarr の開発状況等の情報を常にモニターしており、BD-Zarr の開発方針に反映している。さらに、SSBD:database のメタデータについては、REMBI におけるメタデータをカバーするように更新し、さらに遺伝子関連情報を構造化したかたちで共有できるように改良を行った。顕微鏡の構成についてのメタデータも QUAREP LiMi で推奨されるメタデータを参考に改良を行った。加えて、欧州のバイオイメージングデータの高付加価値データベースである IDR との連携を開始し、遺伝子情報や GO を含むメタデータを利用した相互検索を可能にするための基盤整備やその基

盤を利用した解析シナリオの議論を行った。

3a) バイオイメージングデータに関する新たな情報解析手法の開発のためのデータ基盤の整備

SSBD:database で既に共有されている画像データからいくつかの特徴的なデータを選択し、bioformats2raw ツール (<https://github.com/glencoesoftware/bioformats2raw>) を利用して、OME-Zarr フォーマットへの変換を行った。元画像データに十分なメタデータが付与されていない場合には OME-Zarr に含まれる JSON フォーマットのメタデータを手動で編集する必要があるため、変換したすべての OME-Zarr フォーマットのメタデータを編集して、再利用性の向上をはかった。OME-Zarr フォーマットで格納された画像データの外部公開を行い (<https://ssbd.riken.jp/ssbd-ome-ngff-samples/>)、これらの成果を、国際的なバイオイメージングコミュニティグループと共に、学術論文誌上で発表した (Moore et al., *Histochem Cell Biol*, 2023)。これらのデータ共有は、理化学研究所の共同利用計算システムである HOKUSAI SailingShip (SS) の S3 互換オブジェクトストレージを利用することにより達成した。

3b) バイオイメージングデータに関する主要な画像情報処理ソフトウェアの収集と実行環境の提供

BinderHub を利用してワークフロー実行環境を提供しているサービスを調査し、GitHub 上に置かれた Python や R で記述されたワークフローが実行可能であることを確認した。公開されている GitHub 上のワークフローは、Google Colab や Binder といった外部の実行環境で動作させることは既に可能であり、現在 SSBD:database 側で GitHub 上のワークフローへのリンクを提供するための仕組みを検討している。また、ワークフローを実行する環境として、国立情報学研究所 (NII) が提供しているコード付帯機能 (<https://rcos.nii.ac.jp/service/cs/>) の調査も行い、Gakunin RDM (<https://rdm.nii.ac.jp/>) を利用することにより、GitHub 上に置かれた Python や R で記述されたワークフローを実行可能なことを確認した。一方、画像データの解析ツールとしては、深層学習を基盤としたツールが新たに開発され、利用可能な状況になりつつある。その中で、疎なアノテーションから継続深層学習によりセグメンテーションおよびトラッキングが可能なツール ELEPHANT (<https://elephant-track.github.io>) に注目した。既に多くの実験生物学者から ELEPHANT 利用の要望を受けており、ELEPHANT の実行環境を提供することにした。深層学習に対応した GPU サーバシステムを導入するとともに、その整備とテストを行った。

4a) SSBD データベースの運用の理化学研究所共同利用計算システムへの移行

本年度、当初の予定を繰り上げ、理化学研究所共同利用計算システムの HOKUSAI SS を利用したデータ共有を開始した。SSBD:database で共有されている一部の画像データを OME-Zarr フォーマットに変換し、HOKUSAI SS システム上の S3 互換オブジェクトストレージにてデータ共有を開始した。さらに、OME-Zarr フォーマットで保存されたバイオイメージングデータの高速な転送やアクセスに関して、NTT (日本電信電話株式会社) および理化学研究所情報統合本部と、IOWN (Innovative Optical and Wireless Network) による大規模研究データの利活用促進についての共同研究を開始した。

§4. 成果発表等

(1) 原著論文発表

① 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	3件
	国際(欧文)	6件
未発行論文 (accepted, in press 等)	国内(和文)	0件
	国際(欧文)	0件

② 論文詳細情報

1. Taro Ichimura, Taishi Kakizuka, Yuki Sato, Keiko Itano, Kaoru Seiriki, Hitoshi Hashimoto, Hiroya Itoga, Shuichi Onami, Takeharu Nagai. Volumetric trans-scale imaging of massive quantity of heterogeneous cell populations in centimeter-wide tissue and embryo, *eLife*, vol. 13, RP93633, 2024 (DOI: 10.7554/eLife.93633).
2. Bajcsy, Peter, Bhattiprolu, Sreenivas, Boerner, Katy, Cimini, Beth A., Collinson, Lucy, Ellenberg, Jan, Fiolka, Reto, Giger, Maryellen, Goscinski, Wojtek, Hartley, Matthew, Hotaling, Nathan, Horwitz, Rick, Jug, Florian, Kreshuk, Anna, Lundberg, Emma, Mathur, Aastha, Narayan, Kedar, Onami, Shuichi, Plant, Anne L., Prior, Fred, Swedlow, Jason, Taylor, Adam, Keppler, Antje. Enabling Global Image Data Sharing in the Life Sciences, *arXiv*, 2024 (DOI: 10.48550/arxiv.2401.13023).
3. Nikki Bialy, Frank Alber, Brenda Andrews, Michael Angelo, Brian Beliveau, Lacramioara Bintu, Alistair Boettiger, Ulrike Boehm, Claire M. Brown, Mahmoud Bukar Maina, James J. Chambers, Beth A. Cimini, Kevin Eliceiri, Rachel Errington, Orestis Faklaris, Nathalie Gaudreault, Ronald N. Germain, Wojtek Goscinski, David Grunwald, Michael Halter, Dorit Hanein, John W. Hickey, Judith Lacoste, Alex Laude, Emma Lundberg, Jian Ma, Leonel Malacrida, Josh Moore, Glyn Nelson, Elizabeth Kathleen Neuman, Roland Nitschke, Shuichi Onami, Jaime A. Pimentel, Anne L. Plant, Andrea J. Radtke, Bikash Sabata, Denis Schapiro, Johannes Schöneberg, Jeffrey M. Spraggins, Damiar Sudar, Wouter-Michiël Adrien Maria Vierdag, Niels Volkmann, Carolina Wählby, Siyuan (Steven) Wang, Ziv Yaniv, Caterina Strambio-De-Castillia. Harmonizing the Generation and Pre-publication Stewardship of FAIR Image Data, *arXiv*, 2024 (DOI: 10.48550/arxiv.2401.13022).
4. Josh Moore, Daniela Basurto-Lozada, Sébastien Besson, John Bogovic, Jordão Braganti, Eva M. Brown, Jean-Marie Burel, Xavier Casas Moreno, Gustavo de Medeiros, Erin E. Diel, David Gault, Satrajit S. Ghosh, Ilan Gold, Yaroslav O. Halchenko, Matthew Hartley, Dave Horsfall, Mark S. Keller, Mark Kittisopikul, Gabor Kovacs, Aybüke Küpcü Yoldaş, Koji Kyoda, Alban le Tournoux de la Villegeorges, Tong Li, Prisca Liberali, Dominik Lindner, Melissa Linkert, Joel Lüthi, Jeremy Maitin-Shepard, Trevor Manz, Luca Marconato, Matthew McCormick, Merlin Lange, Khaled Mohamed, William Moore, Nils Norlin, Wei Ouyang, Bugra Özdemir, Giovanni Palla, Constantin Pape, Lucas Pelkmans, Tobias Pietzsch, Stephan Preibisch, Martin Prete, Norman Rzepka, Sameel Samee, Nicholas Schaub, Hythem Sidky, Ahmet Can Solak, David R. Stirling, Jonathan Striebel, Christian Tischer, Daniel Toloudis, Isaac Virshup, Petr Walczysko, Alan M. Watson, Erin Weisbart, Frances Wong, Kevin A. Yamauchi, Omer Bayraktar, Beth A. Cimini, Nils Gehlenborg, Muzlifah Haniffa, Nathan Hotaling, Shuichi Onami, Loic A. Royer, Stephan Saalfeld, Oliver Stegle, Fabian J. Theis & Jason R. Swedlow OME-Zarr: a cloud-optimized bioimaging file format with international community support., *Histochemistry and cell biology*, vol. 160, No. 3, 223-251, 2023 (DOI: 10.1007/s00418-02

3-02209-1).

5. Sugawara, Ko. Training deep learning models for cell image segmentation with sparse annotations, bioRxiv, 2023.06.13.544786, 2023 (DOI: 10.1101/2023.06.13.544786).
6. Dohi, Eisuke, Tatsuya Kushida, Yuki Yamagata, Terue Takatsuki, Jaemoon Shin, Hiroshi Masuya, Thomas Liener, Robert Hoehndorf. BioHackJP 2023 Report R1:Improving phenotype ontology interoperability, BioHackrXiv, 2024 (DOI: <https://doi.org/10.37044/osf.io/d27fw>).

(2) その他の著作物(総説、書籍など)

1. 市村垂生, 糸賀裕弥, 垣塚大志, 大浪修一, 永井健治. 細胞画像ビッグデータによるバイオ DX - イメージング装置と情報プラットフォーム -, 光アライアンス, vol. 34, No. 11, 48-52, 2023.
2. 市村垂生, 糸賀裕弥, 垣塚太志, 大浪修一, 永井健治. 巨大データを生み出すトランススケールスコープとその取扱いを可能にする情報プラットフォーム, 光技術コンタクト, vol. 61, No. 4, 43-50, 2023.
3. 京田耕司, 大浪修一. ライトシート顕微鏡データのとり扱いと共有, 実験医学別冊 最強のステップ UP シリーズ ライトシート顕微鏡実践ガイド組織透明化&ライブイメージング洲崎悦生編, (東京: 羊土社), 163-169, 2023.

(3) 国際学会および国内学会発表

① 概要

種別	国内外	件数
招待講演	国内	5 件
	国際	4 件
口頭発表	国内	4 件
	国際	2 件
ポスター発表	国内	9 件
	国際	3 件

② 招待講演

〈国内〉

1. 大浪修一: バイオイメージングの威力を最大化する情報科学. 第 252 回原医研セミナー/第 16 回放射線災害・医科学研究 機構・拠点研究推進ミーティング, オンライン, 2024 年 1 月 23 日.
2. 大浪修一: Data-driven life sciences using large-scale bioimage data and beyond!. 立命館大学大学院生命科学研究科 ゲノム情報学特論(A), オンライン, 2024 年 1 月 19 日.
3. 大浪修一: Global Image Data Ecosystem, and CREST Multicell imaging data management and sharing. CREST 多細胞領域 第 5 回領域会議, 浜松, 2023 年 12 月 19-21 日.
4. Wen, C., Kimura, K., Onami, S.: Developing an Accurate Method for Tracking Signal Dynamics in Cell Populations within Deforming Organs. 第 32 回日本バイオイメージング学会 学術集会, 札幌, 2023 年 11 月 3-5 日.
5. 大浪修一: SSBD: バイオイメージングデータのグローバルな共有と生命科学の未来. 第 96 回日本生化学大会, 福岡, 2023 年 10 月 31 日-11 月 2 日.

〈国際〉

1. Onami, S.: SSBD: Systems Science of Biological Dynamics. ABIS-GBI 2023 course - Image data: image analysis, data management and reuse, Okazaki and Kobe, Japan, Jul 5-7, 2023.
2. Sugawara, K.: ELEPHANT: tracking cells in 2D/3D+t using interactive deep learning. NEUBIAS Pasteur course on Bioimage Analysis, Paris, France, May 31, 2023.
3. Sugawara, K., Averof, M.: Cell image segmentation with deep learning using sparse annotations. 22nd International European Light Microscopy Initiative Meeting. Noordwijkerhout, The Netherlands, Jun 6-9, 2023.
4. Sugawara, K., Averof, M.: Deep learning with sparse annotations in cell image segmentation. The 5th NEUBIAS Conference. Porto, Portugal, May 11-12, 2023.

③ 口頭講演

〈国内〉

1. 菅原皓: ELEPHANT: an image analysis platform for 2D/3D cell tracking using incremental deep learning. CREST 多細胞 Rising Star Webinar, オンライン, 2023 年 12 月 13 日.
2. Kyoda, K., Itoga, H., Wang, F., Yamamoto, H., Miranda-Miranda, M., Yamagata, Y., Tohsato, Y., Onami, S.: SSBD:repository/SSBD:database: A global sharing of bioimaging data. The 56th Annual Meeting of the Japanese Society of Developmental Biologists, Sendai, Jul 22-25, 2023.
3. 京田耕司, 糸賀裕弥, 王放放, 山縣友紀, ミランダミゲル, 山本春菜, 遠里由佳子, 大浪修一: SSBD: バイオイメージングデータのグローバルなデータ共有. 第 75 回日本細胞生物学会大会, 奈良, 2023 年 6 月 28 日-30 日.
4. 山縣友紀, 大浪修一, 榎屋啓志: 老化制御を目指した細胞老化オントロジーモデリング. 2023 年度 人工知能学会全国大会 (第 37 回), 熊本, 2023 年 6 月 6-9 日.

〈国際〉

1. Sugawara, K.: Training deep learning models for cell image analysis with sparse annotations. CNRS RtmFm Webinar. Online, Dec 8, 2023.
2. Kyoda, K., Itoga, H., Wang, F., Yamamoto, H., Miranda-Miranda, M., Yamagata, Y., Tohsato, Y., Onami, S.: SSBD:repository/SSBD:database: A global sharing of bioimaging data. The 20th International Microscopy Congress, Busan, Sep 10-15, 2023.

④ ポスター発表

〈国内〉

1. 糸賀裕弥, 藤澤絵美, 山本春菜, Miranda-Miranda Miguel, 山縣友紀, 京田耕司, 遠里由佳子, 大浪修一: SSBD: Global sharing of bioimaging data, 第 27 回異分野交流の夕べ, 2024 年 2 月 22 日.
2. 菅原皓, 大浪修一.: ELEPHANT: an image analysis tool for 2D/3D cell tracking using incrementally trainable deep learning. 第 27 回 異分野交流の夕べ. 神戸, 2024 年 2 月 22 日.
3. 京田耕司, 糸賀裕弥, 大浪修一.: Next generation file formats of bioimaging and biosystems dynamics data. 第 27 回 異分野交流の夕べ. 神戸, 2024 年 2 月 22 日.
4. Miranda-Miranda, M., Kyoda, K., Itoga, H., Onami, S.: BD5Viewers: Visualization tools on quantitative biological dynamics in the SSBD. 第 46 回日本分子生物学会年会, 神戸, 2023 年

12月6-8日.

5. 櫛田達矢, 臼田大輝, 高田豊行, 山縣友紀, 榎屋啓志: ヒト哺乳類表現型オントロジーマッピングを活用したヒト疾患モデル生物の推定方法の検討. 第46回日本分子生物学会年会, 神戸, 2023年12月6-8日.
6. Kyoda, K., Itoga, H., Wang, F., Miranda-Miranda, M., Yamamoto, H., Yamagata, Y., Tohsato, Y., Onami, S.: SSBD:repository/SSBD:database: バイオイメージングデータのグローバルな共有. 第61回日本生物物理学会年会, 名古屋, 2023年11月14-16日.
7. 糸賀裕弥, 王放放, 藤澤絵美, 山本春菜, Miranda Miranda Miguel, 山縣友紀, 京田耕司, 遠里由佳子, 大浪修一: SSBD: バイオイメージングデータのグローバルなデータ共有, JST NBDC 統合化推進プログラム2023年度研究交流会, 東京都江東区, 2023年10月5日.
8. 京田耕司, 糸賀裕弥, ミランダミゲル, 大浪修一: OME-Zarr/BD-Zarr: バイオイメージング関連データの次世代フォーマットの開発, JST NBDC 統合化推進プログラム2023年度研究交流会, 東京都江東区, 2023年10月5日.
9. Miranda Miranda Miguel, 糸賀裕弥, 京田耕司, 山縣友紀, 王放放, 藤澤絵美, 山本春菜, 遠里由佳子, 大浪修一: SSBD: A database of images and data on biological dynamics. An infrastructure to preserve and share science resources. 線虫研究の未来を創る会2023, 神戸, 2023年8月17-18日.

〈国際〉

1. Itoga, H., Fujisawa, E., Yamamoto, H., Miranda-Miranda, M., Yamagata, Y., Kyoda, K., Tohsato, Y., Onami, S.: SSBD: Global sharing of bioimaging data. RIKEN BDR Symposium 2024 "Time Across Scales: Development, Homeostasis and Aging", Kobe, March 4-6, 2024.
2. Sugawara, S., Onami, S.: ELEPHANT: an image analysis tool for 2D/3D cell tracking using incrementally trainable deep learning. RIKEN BDR Symposium 2024 "Time Across Scales: Development, Homeostasis and Aging", Kobe, March 4-6, 2024.
3. Kyoda, K., Itoga, H., Wang, F., Yamamoto, H., Miranda-Miranda, M., Yamagata, Y., Tohsato, Y., Onami, S.: SSBD:repository/SSBD:database: A global sharing of bioimaging data. The 56th Annual Meeting of the Japanese Society of Developmental Biologists, Sendai, Jul 22-25, 2023.

(4) 知的財産権の出願 (国内の出願件数のみ公開)

出願件数

種別		件数
特許出願	国内	0件

(5) 受賞・報道等

① 受賞

1. *2023年度人工知能学会全国大会(第37回)全国大会優秀賞, 山縣友紀, 大浪修一, 榎屋啓志. 2023年7月19日.

② メディア報道

1. Euro BioImaging プレスリリース” foundingGIDE – Laying the foundations of a Global Im

age Data Ecosystem”, <https://www.eurobioimaging.eu/news/foundingide--laying-the-foundations-of-a-global-image-data-ecosystem/>, 2024年2月14日

2. Nature, How open-source software could finally get the world’s microscopes speaking the same language, 2023年10月2日. <https://www.nature.com/articles/d41586-023-03064-9>
3. NBDC ブログ”生体組織の立体モデルをぐりぐり動かせる・時間変化を観察できるビューア「QtBD5Viewer」(開発者インタビュー)”公開. 2023年7月7日. <https://biosciencedbc.jp/blog/20230707-01.html>
4. Marx, V. To share is to be a scientist. Nat Methods 20, 984–989. <https://doi.org/10.1038/s41592-023-01927-7>.
5. 名古屋市立大学 プレスリリース” Seg2Link:三次元立体画像中の細胞の形態を人工知能と共同して解析するためのソフトウェアの開発”, <https://www.nagoya-cu.ac.jp/press-news/202305221800/>, 2023年5月22日.

③ その他の成果発表

1. QtBD5Viewer. ソースコード公開, 2023年5月18日.

§5. 主要なデータベースの利活用状況

(1) アクセス数

① 実績

表 1 研究開発対象の主要なデータベースの利用状況 (月平均)

DB名	種別	2022年度	2023年度
SSBD データベース	訪問者数	1,337	1,806
	訪問数	2,174	3,146
	閲覧ページ数	92,614	495,686

② 分析

- ・ 訪問者数、訪問数、閲覧ページ数ともに増加を続けている。
- ・ 2023年6月末から7月初めにかけて、Global BioImaging および ABiS によるバイオイメージに関するトレーニングコースが基礎生物学研究所(岡崎市)および理化学研究所生命機能科学研究センター(神戸市)で実施され、SSBD とともに Image Data Resource/IDR および BioImage Archive/BIA と連携したコースが実施されたことから、その期間においてアクセス数の総計で平時の 5 倍以上、訪問者数および訪問数において月平均で 2 倍以上の実績が得られている。
- ・ SSBD データベースの高機能化にともない、BD5 フォーマットで提供される生命科学定量データおよび OME-Zarr フォーマットで提供される顕微鏡イメージデータへのアクセスが増加している。これらのデータは同一の訪問者により数多くのアクセスをともなうことから、閲覧ページ数およびアクセス数の総計も増加している。

(2) データベースの利用状況を示すアクセス数以外の指標

1. 理化学研究所脳神経科学研究センターシステム分子行動研究チームの梶山十和子研究員(現・東北大学大学院生命科学研究所分子行動分野助教)と共同研究を実施し、SSBD:repositoy における ssbd-repos-000266 を基盤として、新たな高付加価値データベース Adult Zebrafish Brain Gene Expression Database (<http://ssbd.riken.jp/azebex/>)を開発し公開した。この結果をプレプリントとして公開し、

An atlas and database of neuropeptide gene expression in the adult zebrafish forebrain
Towako Hiraki-Kajiyama, Nobuhiko Miyasaka, Reiko Ando, Noriko Wakisaka, Hiroya Itoga, Shuichi Onami, Yoshihiro Yoshihara bioRxiv 2023.03.29.534505; doi: <https://doi.org/10.1101/2023.03.29.534505>

以下の論文として再録された

Towako Hiraki-Kajiyama, Nobuhiko Miyasaka, Reiko Ando, Noriko Wakisaka, Hiroya Itoga, Shuichi Onami, Yoshihiro Yoshihara: An atlas and database of neuropeptide gene expression in the adult zebrafish forebrain, The Journal of Comparative Neurology, <https://doi.org/10.1002/cne.25619>

- 立命館大学の遠里由佳子教授らは、SSBD:repository における ssbd-repos-000140 および SSBD:database における 140-Hirata-MouseBrainSlice にて公開されたマウスの脳スライス画像を用いて、機械学習による解析を行った。その結果を以下の論文として発表した

Yuki Shimojo, Kazuki Suehara, Tatsumi Hirata, Yukako Tohsato: Segmentation of mouse brain slices with unsupervised domain adaptation considering cross-sectional locations, IPSJ Transactions on Bioinformatics, 17, 33-39 (2024/05/01) <https://doi.org/10.2197/ipsjtbio.17.33>

- 立命館大学の遠里由佳子教授らは、SSBD:repository における ssbd-repos-000140 および SSBD:database における 140-Hirata-MouseBrainSlice にて公開されたマウスの脳スライス画像を用いて、機械学習による解析を行った。その結果を以下の国際会議録として発表した

Hiroto Kawabata, Yuki Shimojo, Tatsumi Hirata, Yukako Tohsato, Large-scale image processing and three-dimensional reconstruction of mouse brains with neurogenic-tagged neurons, Proceedings of the 13th International Conference on Biomedical Engineering and Technology (ICBET 2023), Tokyo, Japan, 1-7, <https://dl.acm.org/doi/10.1145/3620679.3620680> (2023/06/15-18, 2023/12/19).

(3) データベースの利活用により得られた研究成果(生命科学研究への波及効果)

- 立命館大学の遠里由佳子教授らは、SSBD:repository における ssbd-repos-000140 および SSBD:database における 140-Hirata-MouseBrainSlice にて公開されたマウスの脳スライス画像を用いて、機械学習による解析を行った。その結果を以下の論文として発表した

Yuki Shimojo, Kazuki Suehara, Tatsumi Hirata, Yukako Tohsato: Segmentation of mouse brain slices with unsupervised domain adaptation considering cross-sectional locations, IPSJ Transactions on Bioinformatics, 17, 33-39 (2024/05/01) <https://doi.org/10.2197/ipsjtbio.17.33>

- 立命館大学の遠里由佳子教授らは、SSBD:repository における ssbd-repos-000140 および SSBD:database における 140-Hirata-MouseBrainSlice にて公開されたマウスの脳スライス画像を用いて、機械学習による解析を行った。その結果を以下の国際会議録として発表した

Hiroto Kawabata, Yuki Shimojo, Tatsumi Hirata, Yukako Tohsato, Large-scale image processing and three-dimensional reconstruction of mouse brains with neurogenic-tagged neurons, Proceedings of the 13th International Conference on Biomedical Engineering and Technology (ICBET 2023), Tokyo, Japan, 1-7, <https://dl.acm.org/doi/10.1145/3620679.3620680> (2023/06/15-18, 2023/12/19).

(4) データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)

特になし。

§6. 研究開発期間中に主催した活動(ワークショップ等)

(1) 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2023年 4月5日	チーム内ミーティング(非公開)	非公開	5人	研究進捗報告のためのミーティング
2023年 4月10日	チーム内ミーティング(非公開)	非公開	7人	研究進捗報告のためのミーティング
2023年 4月24日	チーム内ミーティング(非公開)	非公開	4人	研究進捗報告のためのミーティング
2023年 4月28日	チーム内ミーティング(非公開)	非公開	3人	研究進捗報告のためのミーティング
2023年 5月9日	国際コミュニティでのミーティング(非公開)	オンライン	15人	GBI Image Data WG を主催
2023年 5月12日	チーム内ミーティング(非公開)	非公開	3人	研究進捗報告のためのミーティング
2023年 5月17日	チーム内ミーティング(非公開)	非公開	7人	研究進捗報告のためのミーティング
2023年 5月19日	チーム内ミーティング(非公開)	非公開	2人	ELEPHANT tutorial session
2023年 5月23日	国際コミュニティでのミーティング(非公開)	オンライン	10人	GBI Image Data WG を主催
2023年 6月2日	チーム内ミーティング(非公開)	非公開	4人	研究進捗報告のためのミーティング
2023年 6月13日	NBDC とのミーティング(非公開)	非公開	3人	SSBD での DOI 登録について打合せ
2023年 7月10-11日	共同研究打合せ(非公開)	理研神戸 (ハイブリッド)	10人	IDE(OME)と共同研究打合せ
2023年 7月10-11日	共同研究打合せ(非公開)	理研神戸 (ハイブリッド)	10人	BIA(EMBL)と共同研究打合せ
2023年 7月19日	チーム内ミーティング(非公開)	オンライン	9人	研究進捗報告のためのミーティング
2023年 7月21日	文科省視察(非公開)	理研神戸	10人	文科省ライフ課視察において研究紹介
2023年 9月4日	国際コミュニティでのミーティング(非公開)	オンライン	19名	GBI Image Data WG を主催
2023年 10月24日	チーム内ミーティング(非公開)	オンライン	8人	研究進捗報告のためのミーティング
2023年 10月31日	RIKEN-ZEISS Joint Research: Cloud based Image analysis by ZEISS arivis Cloud and RIKEN HOKUSAI(非公開:所内公開)	オンライン	20人	RIKEN-ZEISS Joint Research: Cloud based Image analysis by ZEISS arivis Cloud and RIKEN HOKUSAI
2023年 11月6日	SSBD アドバイサラーボード説明(非公開)	京都大学	2人	農芸化学会へ SSBD アドバイサラーボードについて紹介
2023年 11月7日	チーム内ミーティング(非公開)	オンライン	8人	研究進捗報告のためのミーティング
2023年 11月8日	SSBD アドバイサラーボード説明(非公開)	名古屋大学	2名	日本バイオインフォマティクス学会へ SSBD アドバイサラーボードについて

年月日	名称	場所	参加人数	目的・概要
	開)			紹介
2023年 11月13日	SSBD アドバイサリ ーボード説明(非公 開)	オンライン	3名	日本植物生理学会へ SSBD アドバイ サリーボードについて紹介
2023年 11月15日	SSBD アドバイサリ ーボード説明(非公 開)	オンライン	3名	日本解剖学会へ SSBD アドバイサリ ーボードについて紹介
2023年 11月21日	SSBD アドバイサリ ーボード説明(非公 開)	オンライン	3名	発生生物学会と ABiS へ SSBD アド バイサリーボードについて紹介
2023年 11月21日	チーム内ミーティ ング(非公開)	オンライン	8名	研究進捗報告のためのミーティング
2023年 11月21日	チーム内ミーティ ング(非公開)	オンライン	3名	研究進捗報告のためのミーティング
2023年 11月22日	国際コミュニティでの ミーティング(非公 開)	オンライン	20名	GBI Image Data WG を主催
2023年 11月27日	チーム内ミーティ ング(非公開)	オンライン	3名	研究進捗報告のためのミーティング
2023年 11月28日	所外関係者とのミー ティング(非公開)	オンライン	3名	メタデータアノテーション業務に関する 打合せ
2023年 11月28日	SSBD アドバイサリ ーボード説明(非公 開)	東京大学	2名	日本癌学会へ SSBD アドバイサリ ーボードについて説明
2023年 12月27日	所内ミーティング(非 公開)	理研北京事務 所	3名	中国でのバイオイメージングデータの 利活用や日中交流イベントについて 意見交換
2023年 12月28日	国際連携について 打合せ(非公開)	Institute of Biophysics(北 京)	5名	Li 助教授と、日本と中国のバイオイメ ージングデータ共有の国際連携につ いて情報収集・意見交換
2024年 1月10日	チーム内ミーティ ング(非公開)	オンライン	3名	研究進捗報告のためのミーティング
2024年 1月30日	所外関係者とのミー ティング(非公開)	オンライン	3名	メタデータアノテーション業務に関する 打合せ
2024年 1月31日	所外関係者とのミー ティング(非公開)	オンライン	7名	SSBD プロジェクト用バイオイメージ ングデータ解析基盤システム導入につ いて打合せ
2024年 1月31日	国際コミュニティでの ミーティング(非公 開)	オンライン	10名	QUAREP-LiMi WG7 sub-group3 meeting 席
2024年 2月7日	企業との意見交換 (非公開)	オンライン	3名	ISO 標準案、SSBD について意見交 換会
2024年 2月8日	チーム内ミーティ ング(非公開)	オンライン	3名	研究進捗報告のためのミーティング
2024年 2月13日	所外関係者との打 合せ(非公開)	オンライン	4名	メタデータアノテーション業務に関する 打合せ
2024年 2月13日	公開イベント打合せ (非公開)	オンライン	5名	GBI_EoE2024 打合せ
2024年2 月20日	チーム内ミーティ ング(非公開)	オンライン	7名	研究進捗報告のためのミーティング
2024年 2月20日	国際コミュニティでの ミーティング(非公 開)	オンライン	21名	GBI - Working Group: Image Data Management 出席
2024年 3月8日	チーム内ミーティ ング(SSBD core	オンライン	3名	研究進捗報告のためのミーティング

年月日	名称	場所	参加人数	目的・概要
	Meeting)			
2024年 3月14日	公開イベント打合せ (非公開)	オンライン	5名	GBI_EoE2024 打合せ
2024年 3月19日	所外関係者との打 合せ(非公開)	オンライン	3名	メタデータアノテーション業務に関する 打合せ
2024年 3月26日	チーム内ミーティン グ(DB meeting)	オンライン	8名	研究進捗報告のためのミーティング
2024年 3月27日	国際コミュニティでの ミーティング(非公 開)	オンライン	20名	GBI Image Data WG を主催

(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2023年 7月3-5日	ABiS-GBI 2023 course - Image data: image analysis, data management and reuse	NIBB カンファレンスセンター (岡崎市)、理研 BDR(神戸市)	21人	顕微鏡施設などで画像を取得する研究者や画像解析を行う研究者、顕微鏡施設管理やデータ管理を行う担当者を対象とした、画像解析、データ管理、及び画像データリポジトリなど、画像データの包括的なトピックを学ぶワークショップ
2023年 7月10日	BDR Seminar (所内セミナー)	理研 BDR	30人	Matthew Hartley 博士による講演
2023年 7月25日	DECODE セミナー (所内セミナー)	理研 BDR	20人	菅原研究員による講演
2023年 10月31日	RIKEN-ZEISS Joint Research: Cloud based Image analysis by ZEISS Arivis Cloud and RIKEN HOKUSAI(所内公セミナー)	オンライン	20人	ZEISS 社との共同ですすめているデータ解析について紹介
2023年 11月3日	理化学研究所神戸事業所一般公開	理化学研究所	400	SSBD データベースのデモンストレーションおよび公開データを用いたコンピュータゲームによるアウトリーチ活動
2023年 11月20日	BDR Seminar(所内セミナー)	理研 BDR	30人	野崎慎博士による講演

以上

別紙1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	SSBD:database		豊富なメタデータと可視化・解析基盤と共にデータを共有することでユーザーに対するデータの再利用を容易にする高付加価値データベース。	https://ssbd.riken.jp		維持・発展	データベース等	Tohsato, Y., Ho, K.H.L., Kyoda, K., Onami, S. SSBD: a database of quantitative data of spatiotemporal dynamics of biological phenomena. <i>Bioinformatics</i> 32, 3471-3479 (2016). (DOI: 10.1093/bioinformatics/btw417)
2	SSBD:repository		最小限のメタデータと共にデータを共有することで、データ公開を行うユーザーに対して迅速なデータ公開サービスを提供するデータリポジトリ。	https://ssbd.riken.jp/repository/		維持・発展	データベース等	
3	Biological Dynamics Markup Language/BD5	BDML/BD5	BDMLのスキーマと仕様、BD5の仕様について	https://ssbd.qbic.riken.jp/bdml/		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767 https://doi.org/10.1371/journal.pone.0237468
4	BDML4DViewer		BDML形式で記述された時空間定量データをImageJ上で可視化するためのプラグイン	https://ssbd.qbic.riken.jp/BDML4DViewer/index.html		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767
5	BD5		BD5形式の時空間定量データの入出力および簡単な解析を行うサンプルコード	https://github.com/openssbd/BD5_samples		維持・発展	ツール等	https://doi.org/10.1371/journal.pone.0237468
6	phenochar		BDML形式で記述された時空間定量データから表現型特徴をオフラインで解析するためのツール	https://ssbd.qbic.riken.jp/phenochar/index.html		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767
7	BDML2BD5		BDMLからBD5に変換するコンバーター	https://github.com/openssbd/BDML2BD5		維持・発展	ツール等	https://doi.org/10.1093/bioinformatics/btu767 https://doi.org/10.1371/journal.pone.0237468
8	BD5LINT		BD5形式で記述されたデータファイルの検証を行うツール	https://github.com/openssbd/bd5lint		維持・発展	ツール等	https://doi.org/10.1371/journal.pone.0237468
9	SSBD OME-NGFF Samples		SSBDデータベースで共有している一部データに対するome-zarrファイルセットを公開している	https://ssbd.riken.jp/ssbd-ome-ngff-samples/		維持・発展	データベース等	https://doi.org/10.1007/s00418-023-02209-1
10	BD-Zarr		BD-Zarrの仕様およびI/Oおよび可視化のサンプルコード	https://github.com/openssbd/bdz		維持・発展	ツール等	
11	QTBD5Viewer		BD5ファイルを可視化するQtをベースとしたデスクトップアプリケーション	https://github.com/openssbd/QTBD5Viewer		新規	ツール等	