

SSBDデータベースによる グローバルなバイオイメージングデータの共有

大浪 修一

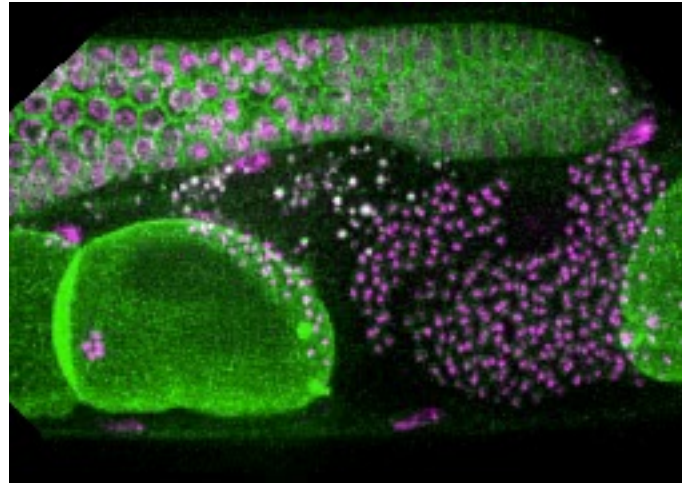
理化学研究所

生命機能科学研究センター 発生動態研究チーム

開拓研究本部 研究デジタルトランスフォーメーション基盤開発チーム

情報統合本部 生命科学データ共有開発ユニット

生命科学におけるイメージングの意義



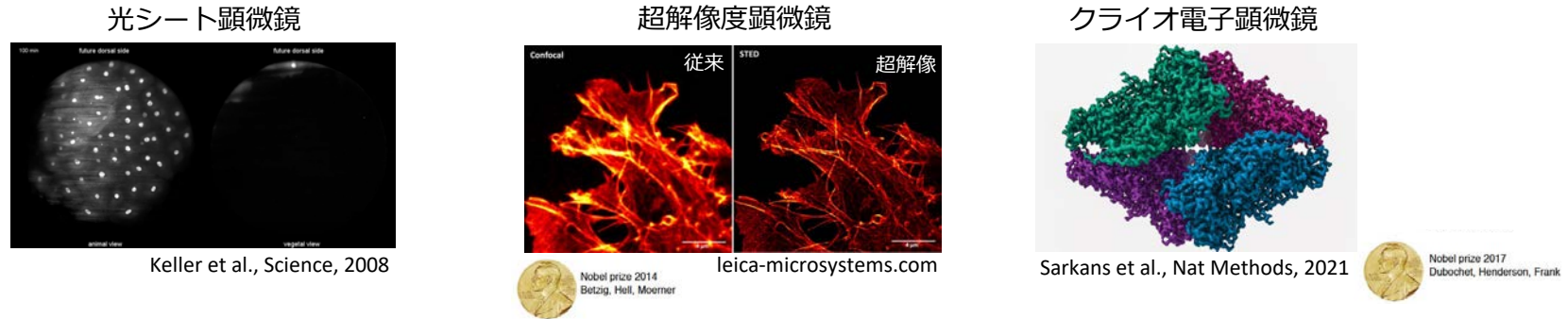
細胞内外の構造や生命分子の空間分布、そしてそれらの動態を詳細にありのままにとらえること

||

生命現象の背景にあるメカニズムを理解するために最も重要な研究手法のひとつ

バイオイメージングの近年の問題

近年の生命科学分野では、最先端のイメージング技術の開発が加速している



↓

最先端のイメージング技術が研究の競争力を決定づける

↓

最先端のイメージング装置は、高度な技術と知識を持つ技術者により手作りで開発される。市販品は極めて高額。

↓

大多数の生命学者は最先端イメージング技術を自らの研究に活用できていない

↓

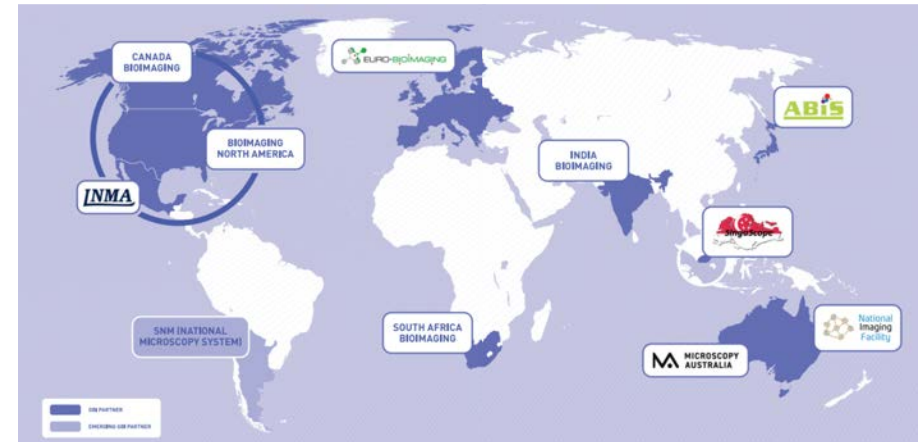
【解決策】

- 最先端のイメージング技術のオープンアクセス化
- 最先端のイメージングデータのオープンアクセス化

GLOBAL BIOIMAGING

growing collaboration

- 生命科学の最先端イメージング技術とデータのオープンアクセスに関連する諸問題を解決する国際連携組織
- 欧州のEuro-BioImagingが全世界のコミュニティに声をかけてボトムアップ的に組織化
- 2015年に欧州の国際共同研究資金（Horizon 2020）の支援で設立
- 2020年よりChan Zuckerberg Initiativeより助成を受ける



<https://www.globalbioimaging.org>

主な活動

- 国際的なリーダーからバイオイメージング施設の運営や研究政策、技術トレンドを学ぶ**年1回の国際ワークショップ（Exchange of Experience）**の開催
- 特定のテーマに関する議論や国際的な協力関係を構築するための**テーマを絞った会議や作業部会**の開催
- バイオイメージング施設の管理職や技術職のスタッフの専門的な能力開発を支援する**研修**の開催
- バイオイメージング施設のスタッフが、海外の先進的な施設から学ぶことを可能にする**スタッフシャドーイング**プログラムの実施

イメージデータ作業部会

共同座長

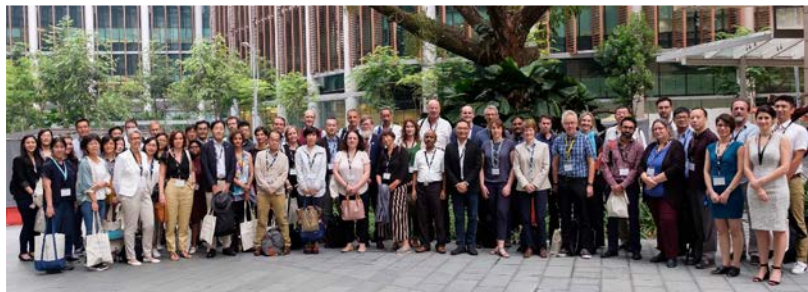


Jason Swedlow
Univ. of Dundee
OME, IDR
-2021



Shuichi Onami
RIKEN BDR
SSBD

- 11カ国、17名の委員。南極を除くすべての大陸より委員が参加。
- バイオイメージデータのオープンサイエンスに関わる諸問題について包括的に議論
 - データフォーマットの標準化
 - 国際的な公共データレポジトリの整備
 - 取得から公開までのデータの管理
 - データ解析情報の管理



Exchange of Experience IV @Singapore 2019



Exchange of Experience VI @Online 2021



A global view of standards for open image data formats and repositories

Imaging technologies are used throughout the life and biomedical sciences to understand mechanisms in biology and diagnosis and therapy in animal and human medicine. We present criteria for globally applicable guidelines for open image data tools and resources for the rapidly developing fields of biological and biomedical imaging.

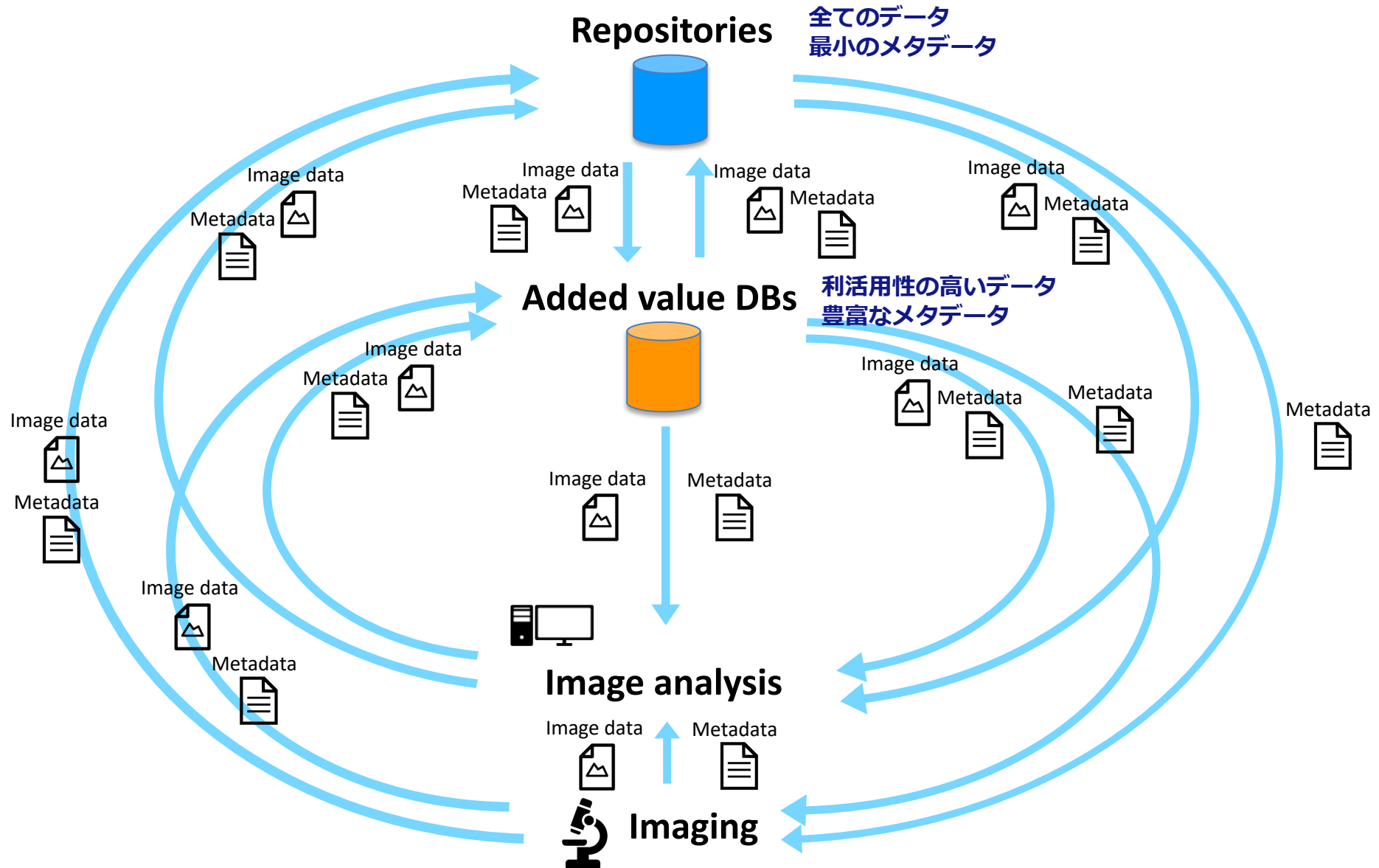
Jason R. Swedlow, Pasi Kankaanpää, Ugis Sarkans, Wojtek Goscinski, Graham Galloway, Leonel Malacrida, Ryan P. Sullivan, Steffen Härtel, Claire M. Brown, Christopher Wood, Antje Keppler, Federica Paina, Ben Loos, Sara Zullino, Dario Livio Longo, Silvio Aime and [Shuichi Onami](#)

Imaging is now used globally as a method of quantitative measurement of biological and biomedical structure, composition and dynamics in the life and biomedical sciences. Imaging technology is rapidly evolving, with new modalities and applications appearing that enable new insights and discoveries^{1,2}. These innovations present challenges at several different but interdependent levels. Sourcing and retaining expert research technology professionals ('imaging scientists'), providing initial and ongoing training in advanced technologies, rapidly disseminating and offering easy access to new innovative methods and applications, publishing reproducible

multiplex tissue imaging, spatial profiling of single-cell transcriptomes, mass spectrometry-based imaging, correlative imaging techniques, molecular imaging, advanced forms of microscopy-based spectroscopy (fluorescence correlation, Raman, hyperspectral) and several others, data complexity and dimensionality are increasing, which makes the need for open, common methods for recording imaging metadata even greater. Moreover, with the establishment and growth of public image data repositories, proposals for common metadata standards are now emerging. It is essential we define the specifications and usability requirements for data standards and repositories so that the global

an important role in defining policy, practice and implementation. Journals have repeatedly contributed to the use of open data standards by requiring that papers submitted for publication adopt domain-specific data deposition standards (for example, deposition of sequence data in ArrayExpress or the European Nucleotide Archive and deposition of structural data in the Protein Data Bank). Funders contribute by conditioning funding awards on the use and adoption of data standards and, where appropriate, the deposition of datasets in open repositories. Finally, Global BioImaging seeks to engage with the community to develop a set of standards and guidelines for open image data. *Nature Methods* 18, 1440-1446 (2021)

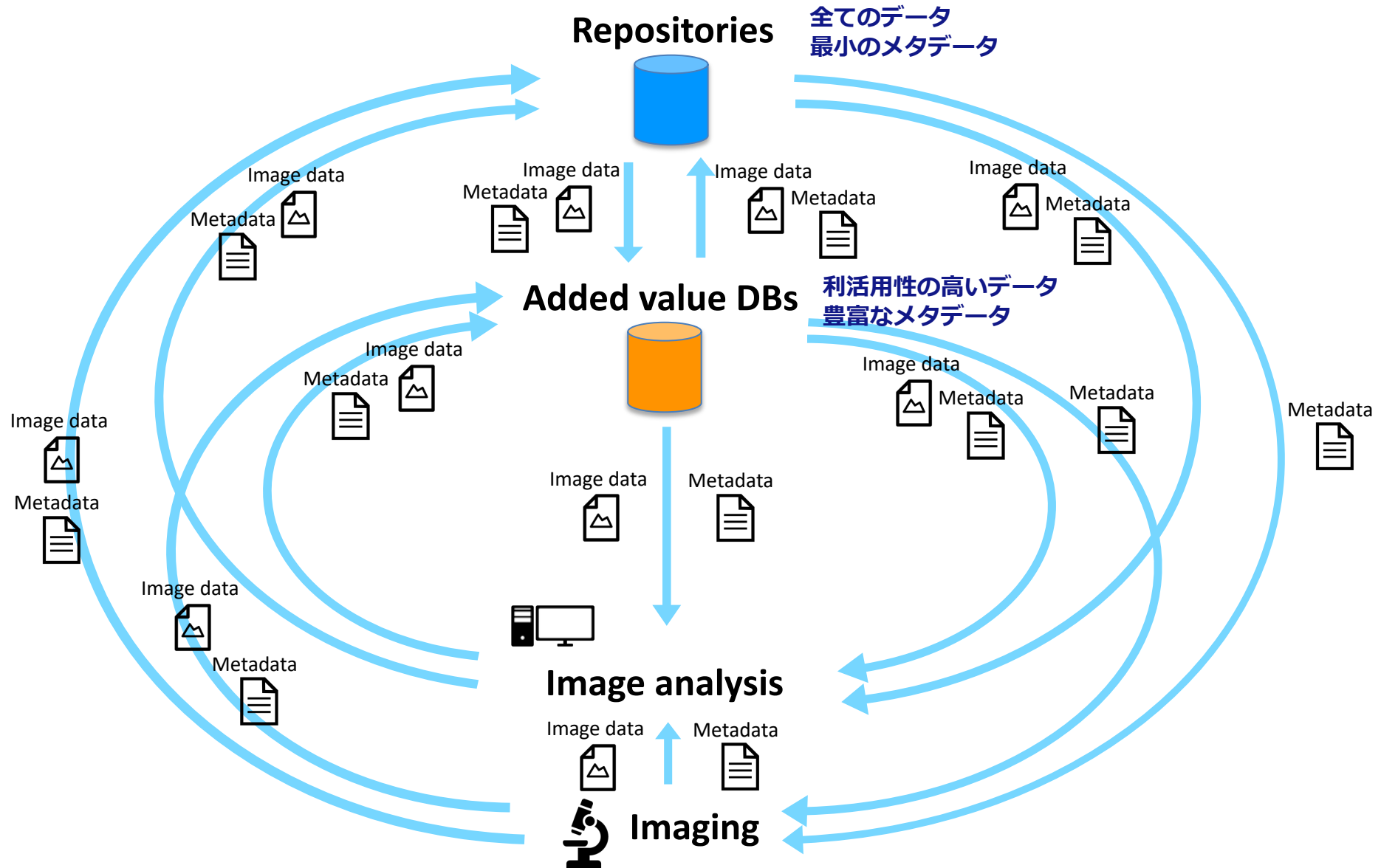
バイオイメージングデータのデータ共有システムのデザイン



データ共有のジレンマ

- FAIR原則 (Wilkinson et al., 2016)
 - **F**indability, **A**ccessibility, **I**nteroperability, **R**eusability
(見つけれれる。アクセスできる。相互運用できる。再利用できる。)
- FAIR原則と現実のジレンマ
 - データは作成後、**できるだけ直ぐに**登録され、見つけれられ、アクセスできるようにしたい。
 - データは適切にキュレーションされ、相互運用可能で再利用可能な形式で保存され、さらなる研究のために利用されるべきである。
 - キュレーション、データ変換には**時間と労力が必要**

バイオイメージングデータのデータ共有システムのデザイン



SSBD:database

SSBD:database

[SSBD:database](#)
[SSBD:repository](#)
[SSBD:OMERO](#)
[Resources](#)
[Tools](#)
[About](#)

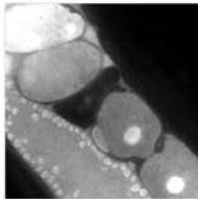
Organism: ex. C. elegans

Introduction

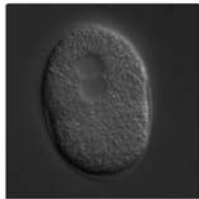
System Science of Biological Dynamics database (SSBD:database) provides a rich set of open resources for analyzing quantitative data and microscopy images of biological objects, such as single-molecule, cell, gene expression nuclei, etc. Quantitative biological data and microscopy images are collected from a variety of species, sources and methods. These include data obtained from both experiment and computational simulation.

Samples

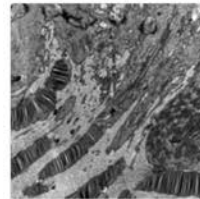
Microscopy Images



Calcium response and shape changes in oocyte of C. elegans



DIC image of nuclear division dynamics in C. elegans embryo



TEM image of retinal tissue from human embryonic stem cells

News

2020-12-24,
SSBD 2020 Update
SSBD 2020 update, the metadata database fully renewed, and 77 image datasets (7GB) and 4 quantitative datasets are added. We are planning next database update in a half year for COVID-19 situation.

2020-08-12,
A new paper on BD5 is published in PLOS ONE
A new paper "BD5: an open HDF5-based data format to represent quantitative biological dynamics data" is now published in PLOS ONE.
<https://doi.org/10.1371/journal.pone.0237468>

[Older news...](#)

Funding



再利用性の高いバイオイメージングデータと生命動態定量データを豊富なメタデータを付与して共有する高付加価値データベース (**Added-Value DB**)

SSBD:repository

SSBD:repository

[SSBD:database](#)
[SSBD:repository](#)
[Resources](#)
[About](#)

Organism: ex. C. elegans

Introduction

System Science of Biological Dynamics repository (SSBD:repository) is an open data archive that stores and publishes bioimaging and biological quantitative datasets that are associated with published or to be published studies. It allow other researchers to access and download those datasets for referene or for further investigations.

Find the dataset from search box above, or see the dataset list in [Resources](#).

SSBD:repository has started operation in 2016, under the life science database integration promotion project of the Japan Science and Technology (JST), National Bioscience Database Center (NBDC). Currently it is funded from RIKEN, JST and Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

For an overview of the SSBD:database/repository, please refer to the paper, Tohsato, Y., Ho, K. H. L., Kyoda, K., and Onami S. (2016) "SSBD: a dataset of quantitative data of spatiotemporal dynamics of biological phenomena." *Bioinformatics*, 32(22): 3471-3479
<https://doi.org/10.1093/bioinformatics/btw417>

Share your data in SSBD:repository

SSBD:repository covers a wide range of biological phenomena, from single molecules to cells / individuals. DOI (Digital Object Identifier) can be assigned to each dataset. Datasets in SSBD:repository are licensed under [CC BY](#), [CC BY-NC-SA](#), etc, upon request of the data contributors.

If you would like to share your image data or quantitative data of biological phenomena, or any

News

2021-02-09,
System maintenance on Feb 9
Due to system maintenance, the 4D visualizer and APIs will not be available on Feb 9, 9:00-16:00. Sorry for the inconvenience.

2020-12-24,
SSBD 2020 Update
SSBD 2020 update, the metadata database fully renewed, and 77 image datasets (7GB) and 4 quantitative datasets are added. We are planning next database update in a half year for COVID-19 situation.

2020-08-12,
A new paper on BD5 is published in PLOS ONE
A new paper "BD5: an open HDF5-based data format to represent quantitative biological dynamics data" is now published in PLOS ONE.
<https://doi.org/10.1371/journal.pone.0237468>

[Older news...](#)

Funding



論文で使用したバイオイメージングデータと生命動態定量データを共有する公共**リポジトリ**

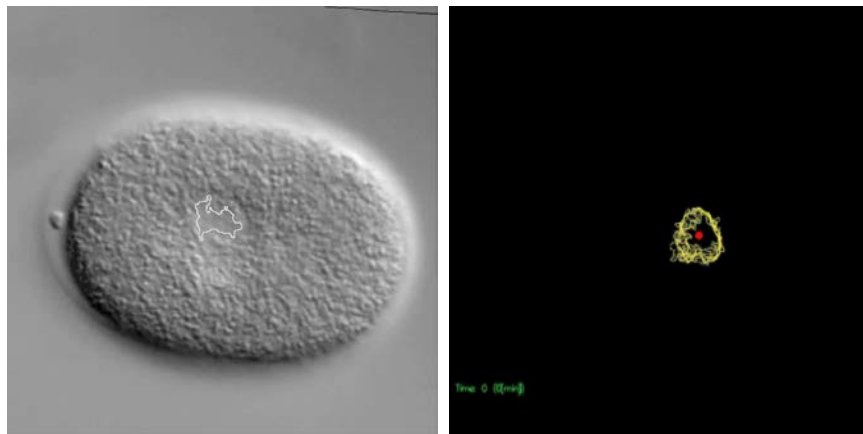
SSBD: バイオイメージングデータと生命動態情報の統合データベース

- 生命動態システム科学の統合データベースとしてJST-NBDCの資金で設立 (2013年)
- 当初は、時空間情報を数値として含む新しい様式の生命科学の研究データとそれを取得するために用いた画像データの共有を行うDB

具体例

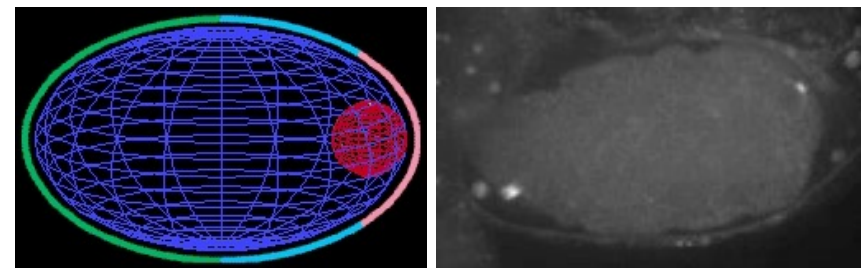
- 様々な摂動条件下の生命動態の**時空間定量計測**データ
- 様々なパラメータ下の生命現象の**時空間動態シミュレーション**結果

線虫胚の細胞核分裂動態計測データ



Hamahashi et al. BMC Bioinformatics 2005

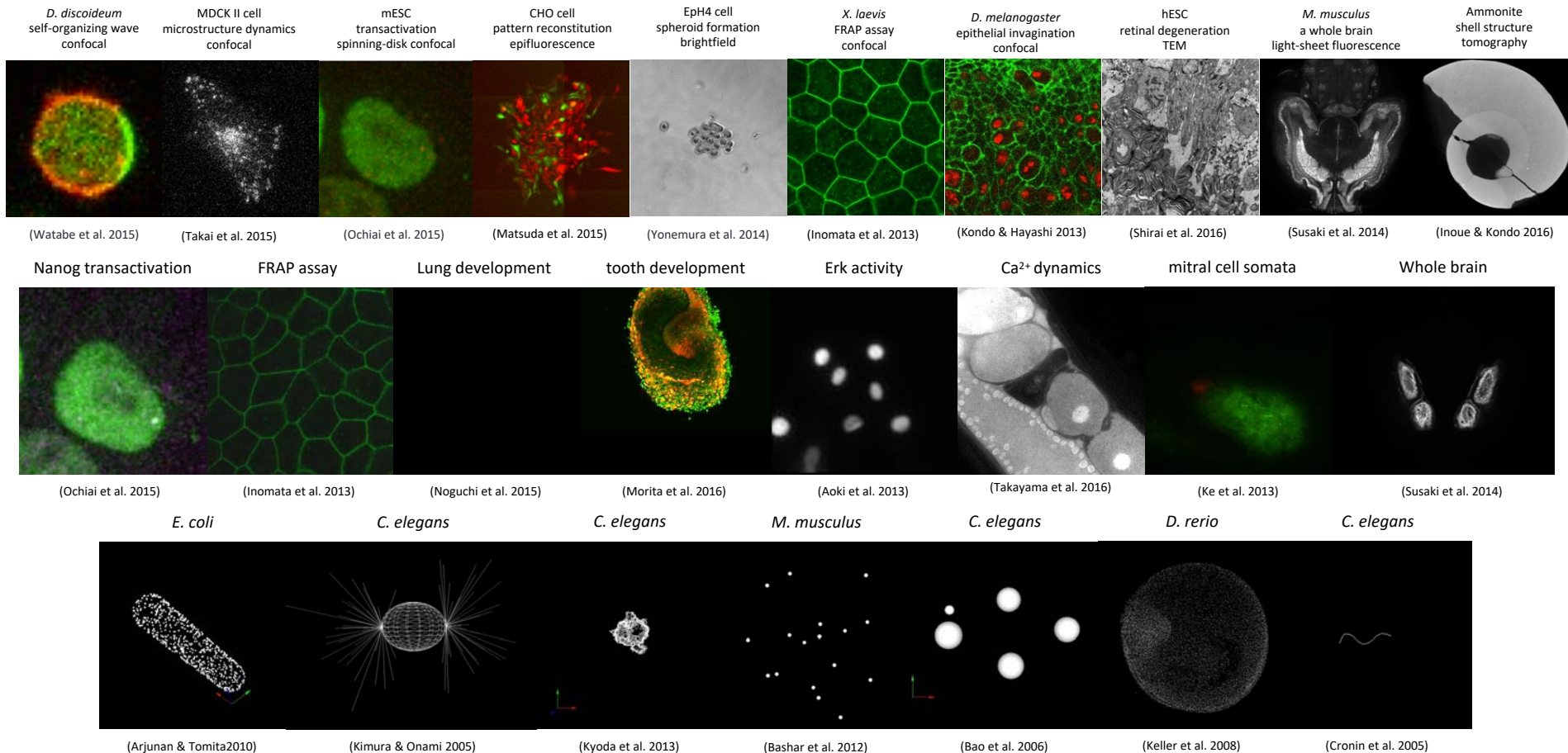
細胞核動態のシミュレーション結果



Kimura & Onami Dev. Cell 2005

SSBD: バイオイメージングデータと生命動態情報の統合データベース

- 最先端技術で取得した生命科学画像データ、網羅的実験の生命科学画像データの共有を開始 (2016年)
- 生命科学画像データのリポジトリサービスを開始 (2016年)
- 高付加価値DB (Added-Value DB) (**SSBD:database**) と公共リポジトリ (**SSBD:repository**) を分離 (2019年)



SSBD:世界を代表するバイオイメージングデータの統合データベース

Work / Technology & tools

A HOME FOR EVERY IMAGING DATA SET

Repositories let researchers store, share and access life-science images – and maybe even extract new findings. **By Amber Dance**

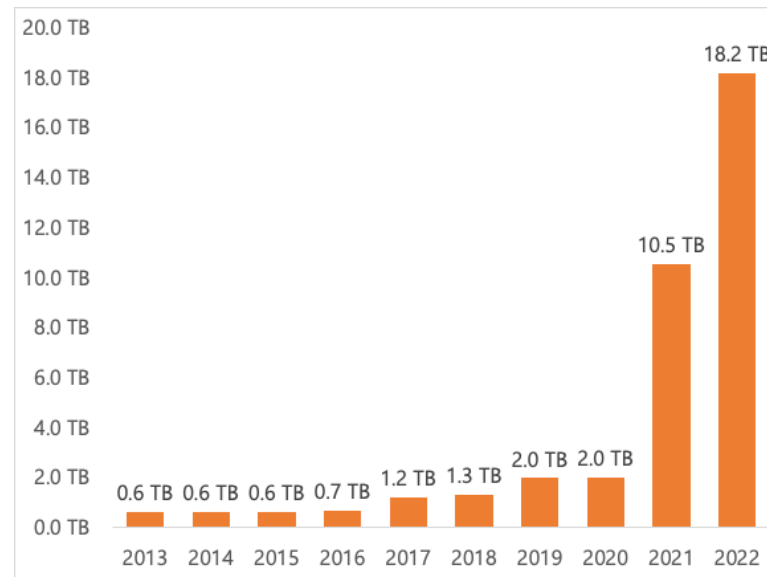
When Sjors Scheres set out to develop a tool to reverse flaws in cryo-electron microscopy images, he needed lots of data on which to test it. So Scheres, a structural biologist at the MRC Laboratory of Molecular Biology (LMB) in Cambridge, UK, turned to the Electron Microscopy Public Image Archive (EMPIAR), a database of raw images. There he downloaded, for free, data collected by the lab of Gabriel Lander, a structural biologist at Scripps Research in La Jolla, California. Using his new technique, Scheres was able to squeeze sharper images from those data, improving the resolution of one structure from 3.1 ångströms to 2.3 ångströms. "That's precisely why we posted the data," says Lander. "We knew some brilliant people out there would be able to improve our processing."

Services such as EMPIAR give researchers a central location in which to store, share and access a rapidly expanding corpus of biological images. "The data aren't just one picture any more," says Joshua Vogelstein, a neurostatistician at Johns Hopkins University in Baltimore, Maryland. Movies, 3D images and microscope-based screening data can take up gigabytes or terabytes of storage, and can't be e-mailed back and forth in the same way as individual TIFF or JPEG files. Moreover, grant agencies and journals increasingly require scientists to make their data available to all, but don't necessarily offer to host them. EMPIAR and its kin fill that gap, and often provide a digital object identifier or other citation so researchers can get credit for their data. "Are you struggling to load your images?" asks Forrest Collman, a neuroscientist at the Allen Institute for Brain Science in Seattle, Washington. "Are you particularly struggling to share?" If so, he says, "looking into this kind of service makes sense for you".

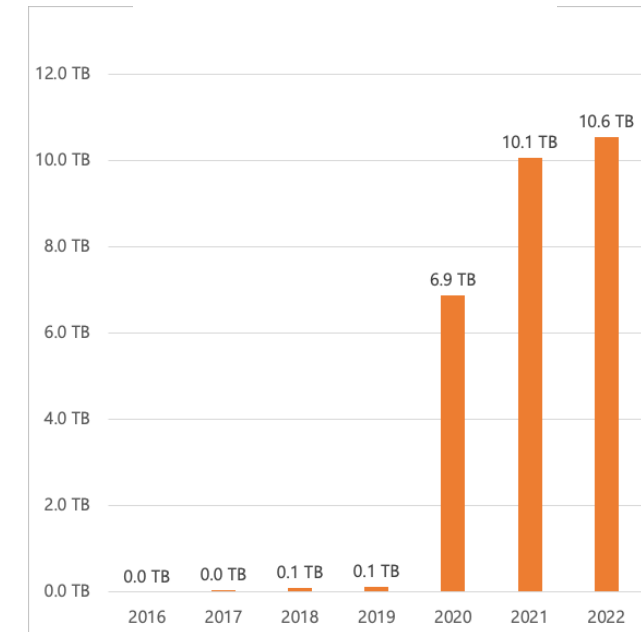
In 2019, when Collman spotted an odd-looking neuron in one of his electron microscopy data sets, it was easy for him to send a colleague a link to that spot in the data repository, rather than a bulky file. She noticed another unique feature, and Collman identified a few similar cells. They might turn out to be a new type of neuron, Collman says.

There are a number of other image warehouses available, among them the Image Data Resource (IDR). Both it and EMPIAR are hosted by the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) in Hinxton, UK. Further options include,

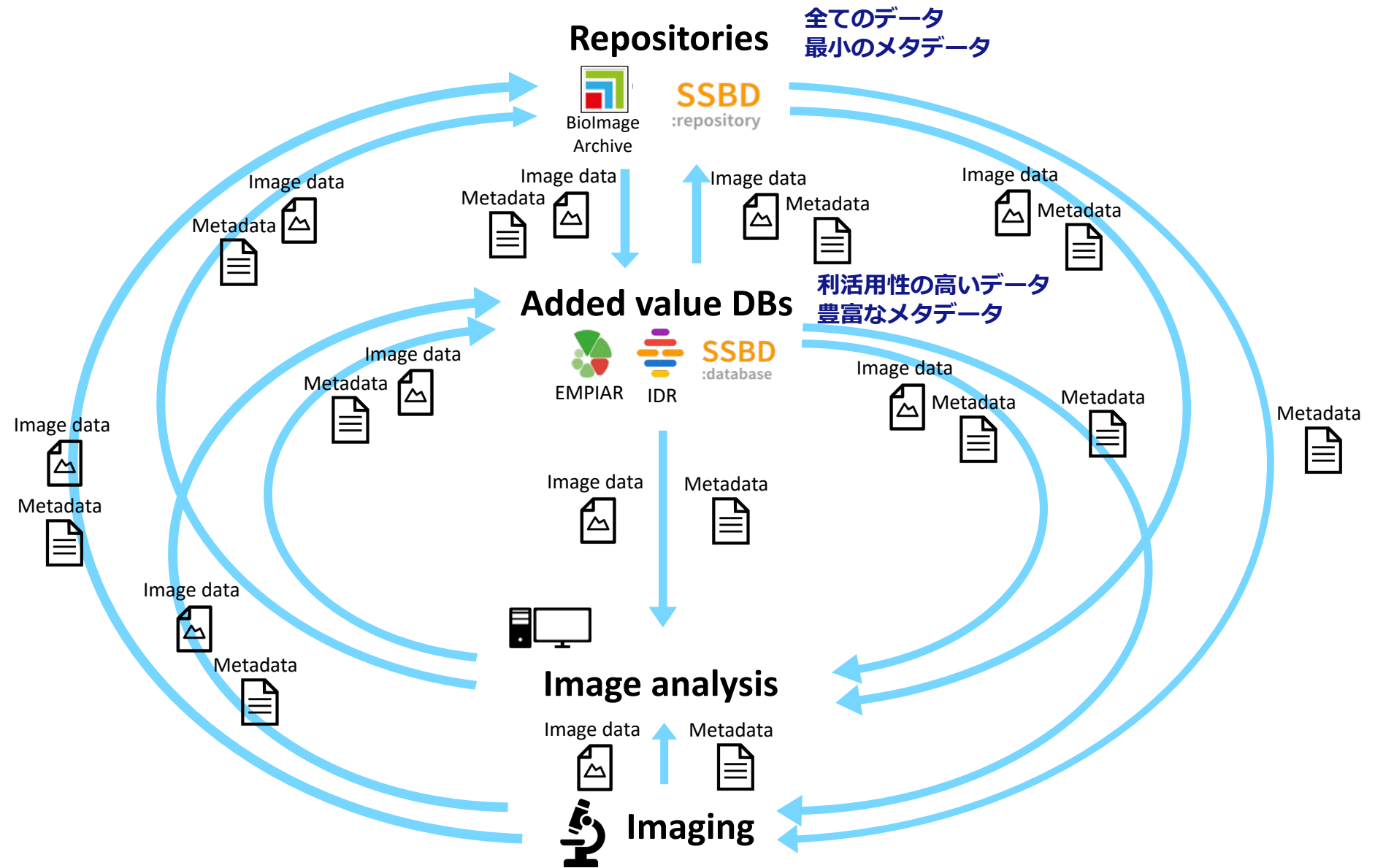
Image data size on **SSBD**:database



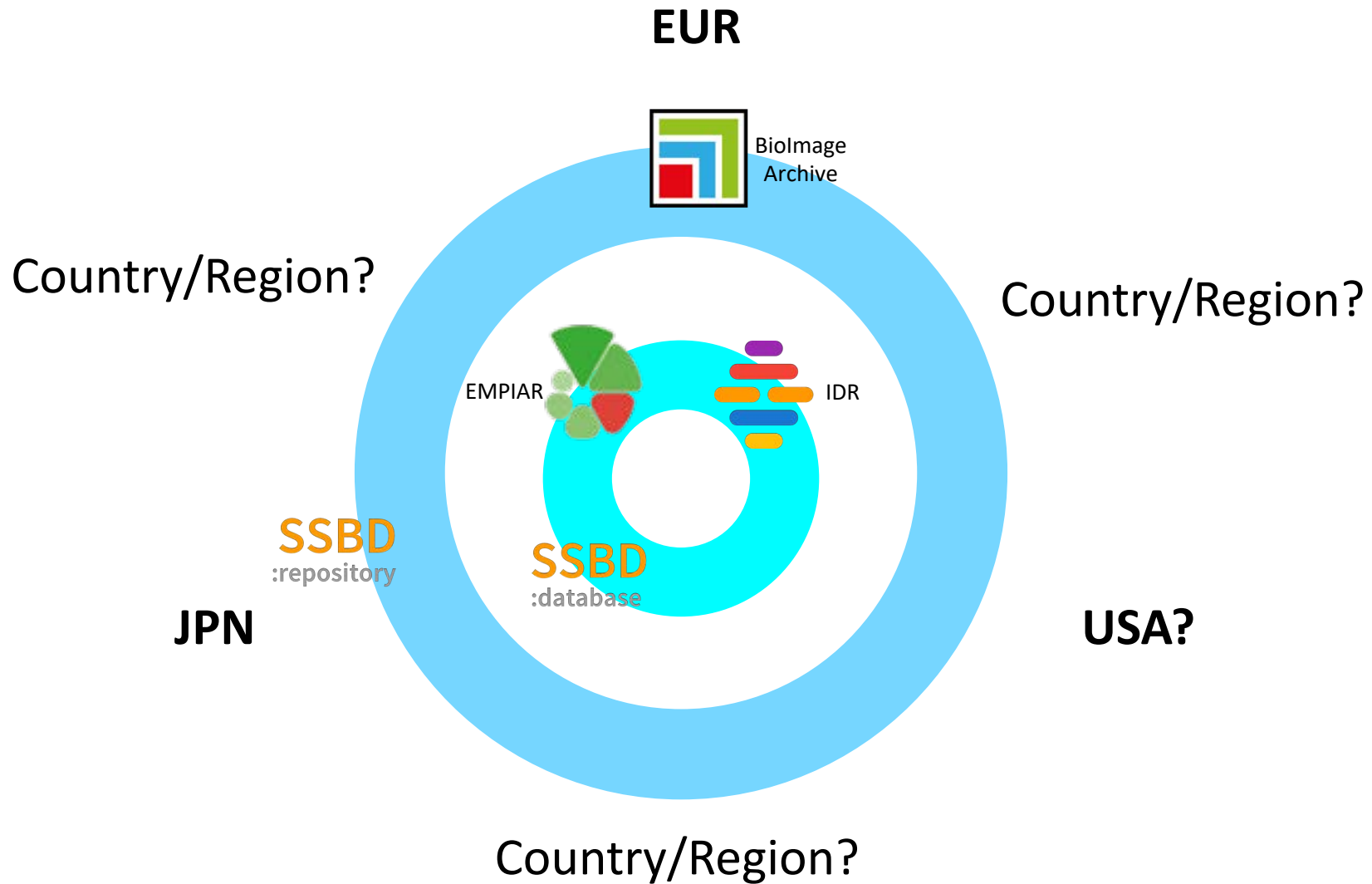
Data size on **SSBD**:repository



日欧のバイオイメージングデータのエコシステム



生命科学画像データのグローバルな共有システム



データを共有する意義

データの共有はデータ解析手法の開発を加速する

ゲノム塩基配列データ、タンパク質立体構造データの集積が、
バイオインフォマティクス技術の発展を可能にした

SSBD:repository へのデータの登録方法

<https://ssbd.riken.jp/repository/share-your-data/>

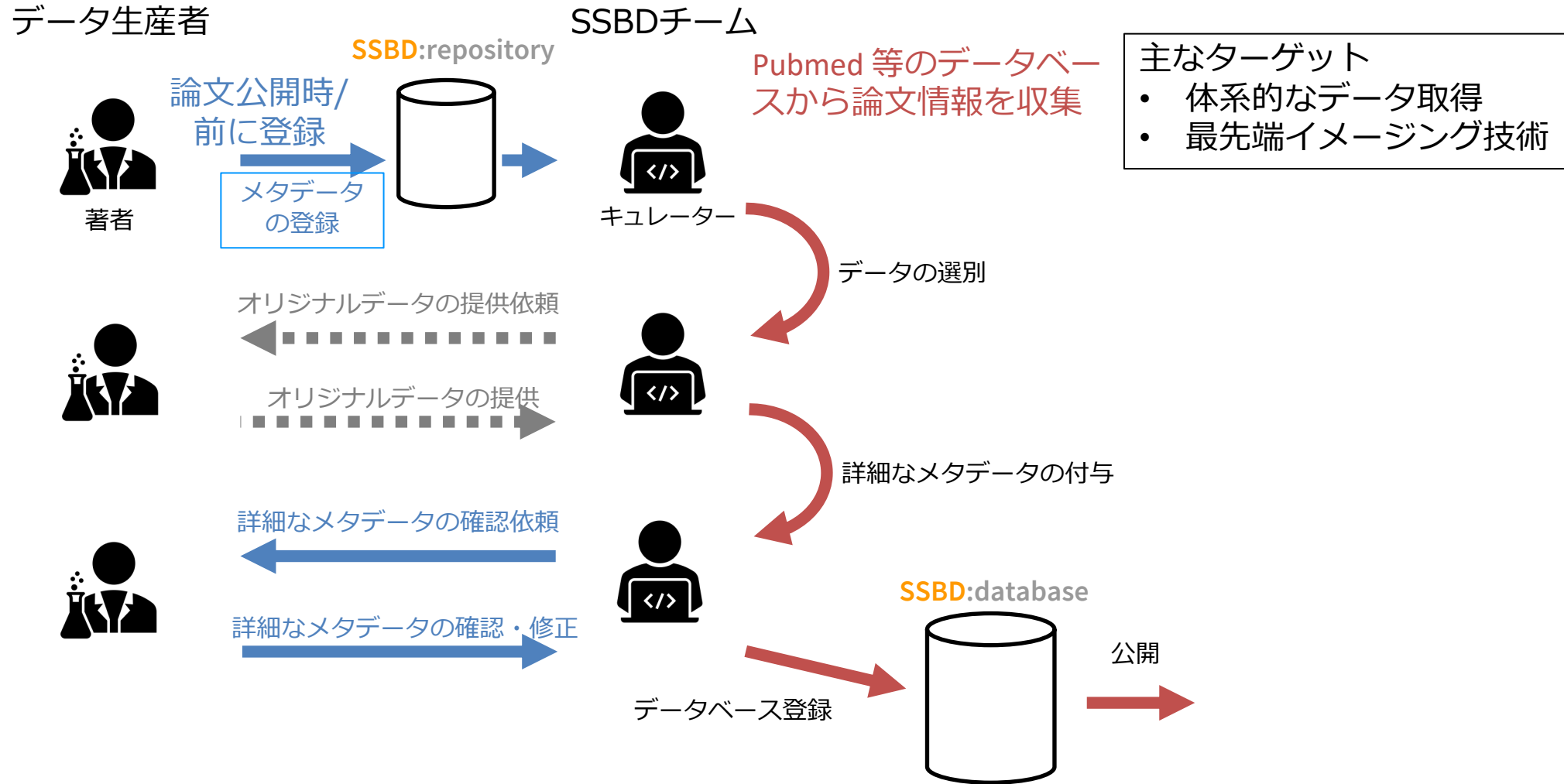
まずは、メールでお問い合わせ下さい

ssbd-repos@ml.riken.jp

お問い合わせのタイミング

- 論文投稿時、リバイス投稿時
 - Reviewerのみがデータを閲覧できるようにする等の対応も可
 - 基本的に全ての論文雑誌のデータ公開の要件と合致
 - 論文中にデータのDOIを表記できる
- △論文採択後
 - データの登録に要する時間に余裕がない

SSBD:database へのデータ登録手順



- データの共有の依頼が来た際にはご協力をお願いします。
- 体系的なデータ取得や最先端イメージングを実施する際は、事前コンタクト！

バイオイメージングデータのグローバルな共有に SSBDデータベースをご利用ください。

<https://ssbd.riken.jp>

SSBD team



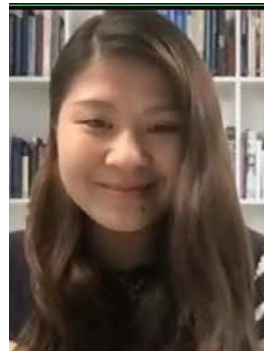
S. Onami



K. Kyoda



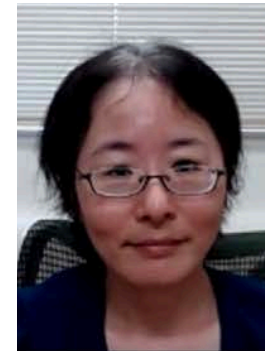
H. Itoga



F. Wang



Y. Yamagata



Y. Tohsato



M. Miranda



H. Yamamoto