

P-15

# SSBD:database and SSBD:repository

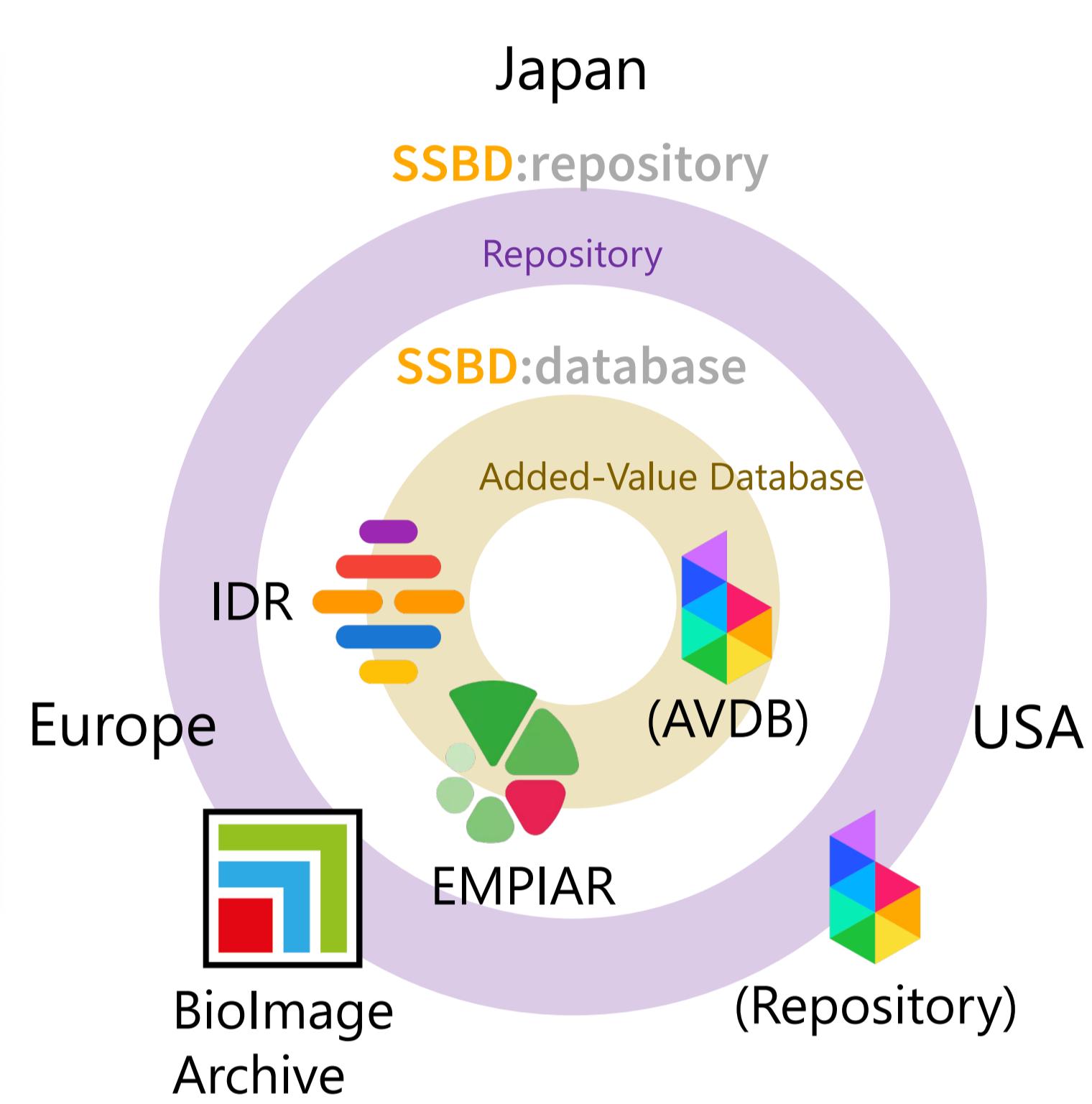
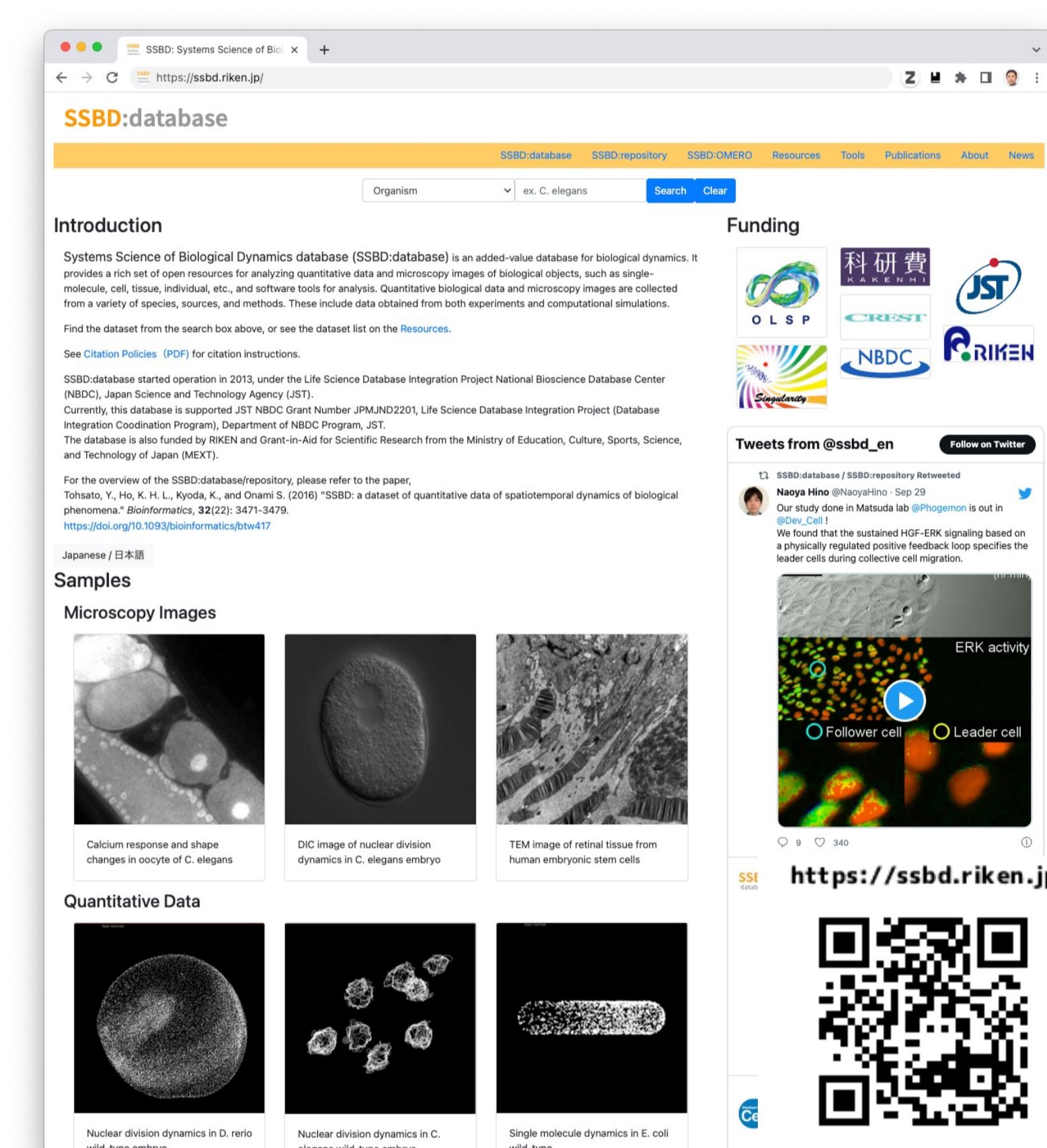
## – Global sharing of bioimaging data –

Hiroya Itoga<sup>1</sup>, Fangfang Wang<sup>1,2</sup>, Yuki Yamagata<sup>1,2</sup>, Koji Kyoda<sup>1</sup>, Yukako Tohsato<sup>1,3</sup>, Shuichi Onami<sup>1,2</sup>

<sup>1</sup> Laboratory for Developmental Dynamics (Onami Lab), RIKEN BDR, Japan  
<sup>2</sup> Life Science Data Sharing Unit, RIKEN Information R&D and Strategy Headquarters, Japan  
<sup>3</sup> Faculty of Information Science and Engineering, Ritsumeikan University, Japan

### Introduction

- SSBD (Systems Science of Biological Dynamics; Tohsato et al. 2016) is an infrastructure for sharing and reusing bioimages and biological dynamics data
- SSBD database is an added-value database (AVDB) for bioimages and biological dynamics data that are highly reusable, taken by state-of-art microscopy, or large-scale data through systematic experiments, with rich metadata
- SSBD repository is a quickly sharing service for all kinds of bioimaging data for published or to-be-published papers, with essential metadata
- Developing a global sharing system among Japan, Europe, and North America, for sharing and reusing bioimages and their metadata

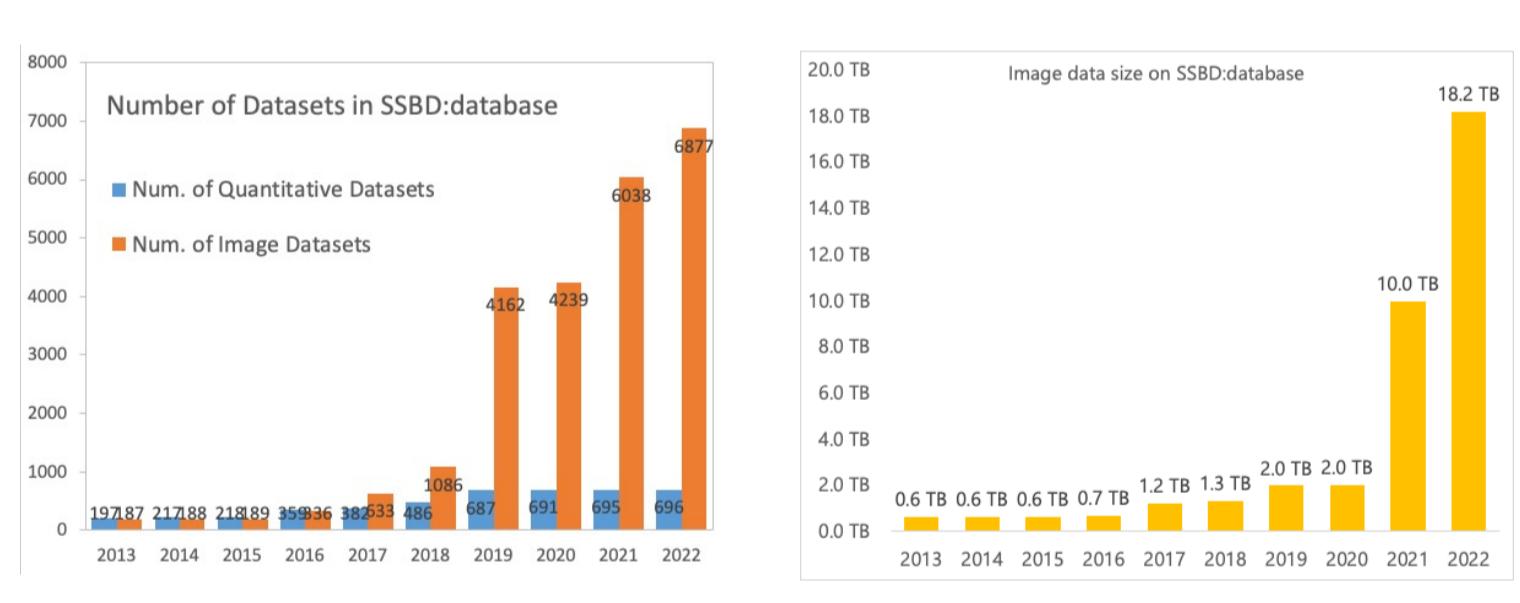


### Resources

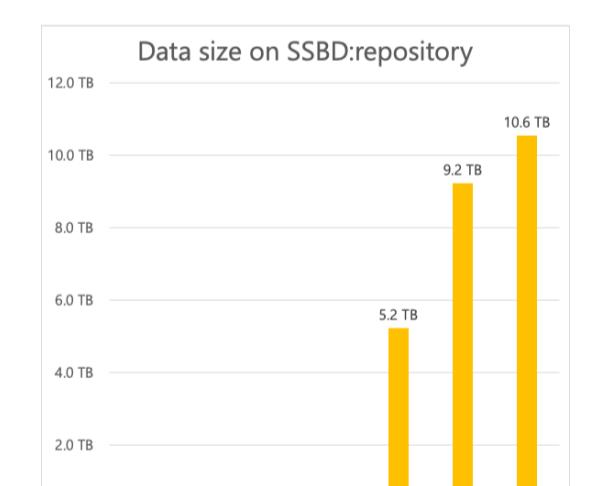
- Bioimaging data
- Biological dynamics data
- Software tools, Workflows

- original, unprocessed data
- analyzed, quantitative (numerical) data

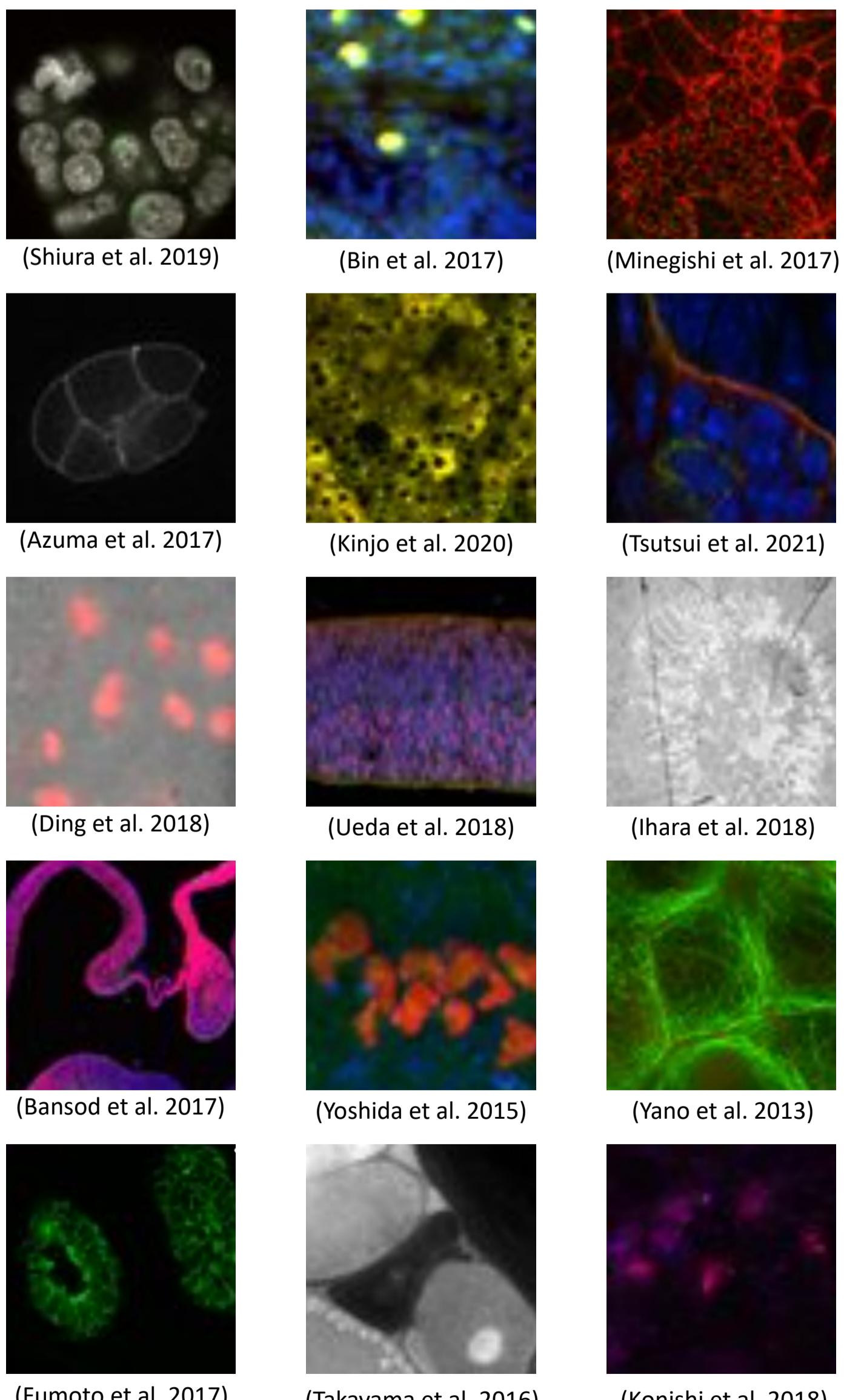
#### Statistics of SSBD:database



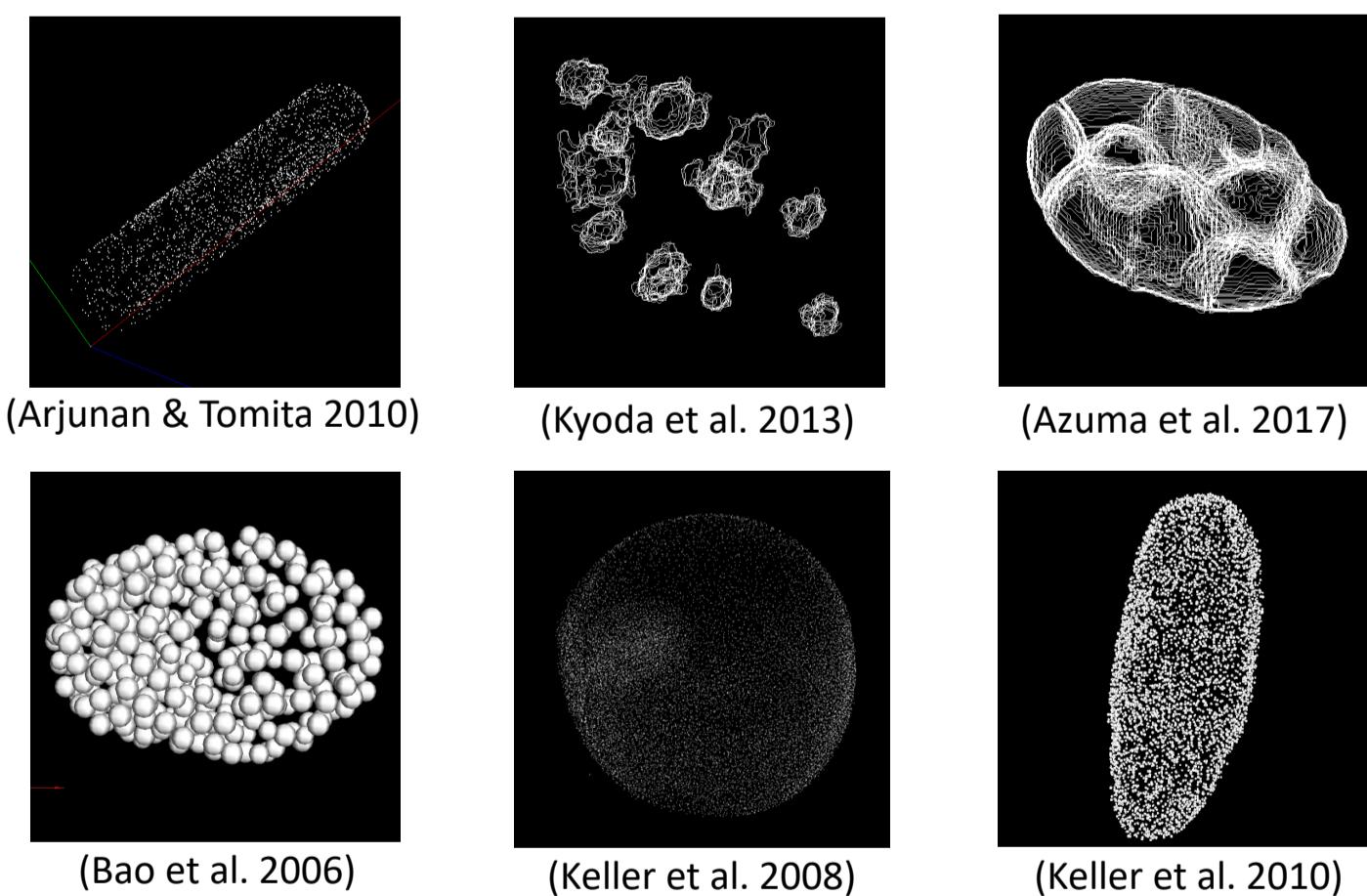
#### Statistics of SSBD:repository



#### Examples of bioimaging data



#### Examples of biological dynamics data



#### Example of workflow

```

A simple SSBD wrapper API to access images
In [0]: source = "https://ssbd.repo/api/getImage"
        BASE_URL = "http://ssbd.repo/api/image"
        file_id = "00000000000000000000000000000000"
        image_id = BASE_URL + "getimage/render_image?file_id=" + file_id + "&format=png"
        resp = requests.get(image_id)
        if resp.status_code == 200:
            print("Success")
            sys.exit()
        else:
            print("Error")
            sys.exit()
        resp.raise_for_status()
        return resp.content

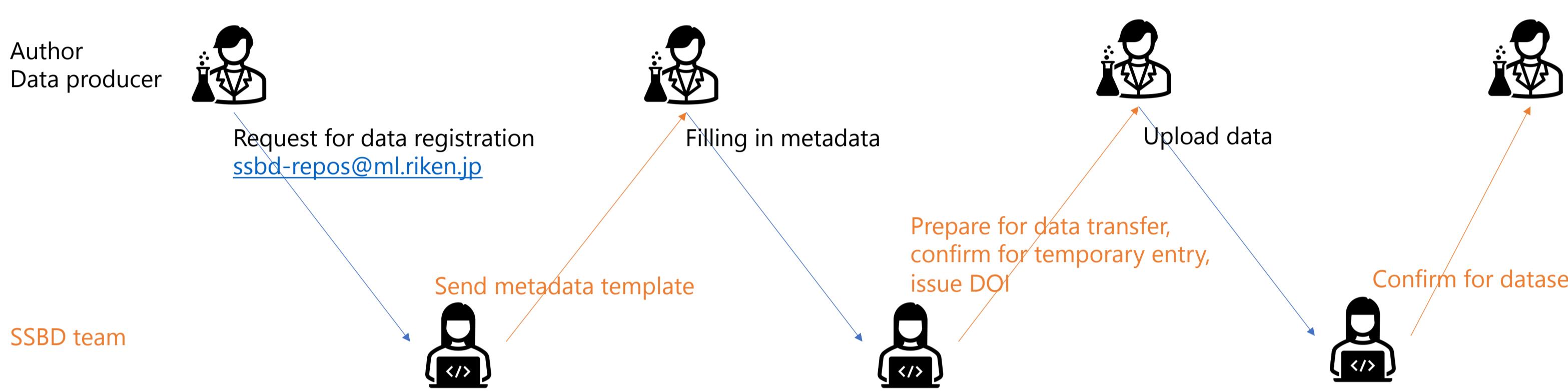
def get_metadata(file_id):
    source = "https://ssbd.repo/api/getImage"
    BASE_URL = "http://ssbd.repo/api/image"
    file_id = "00000000000000000000000000000000"
    image_id = BASE_URL + "getimage/render_image?file_id=" + file_id + "&format=png"
    resp = requests.get(image_id)
    if resp.status_code == 200:
        print("Success")
        sys.exit()
    else:
        print("Error")
        sys.exit()
    resp.raise_for_status()
    return resp.json()

In [1]: print(get_metadata("00000000000000000000000000000000"))
Out[1]: {'file_id': '00000000000000000000000000000000', 'image_id': 'http://ssbd.repo/api/image/AverageImage@0x725e8e1760'}

```

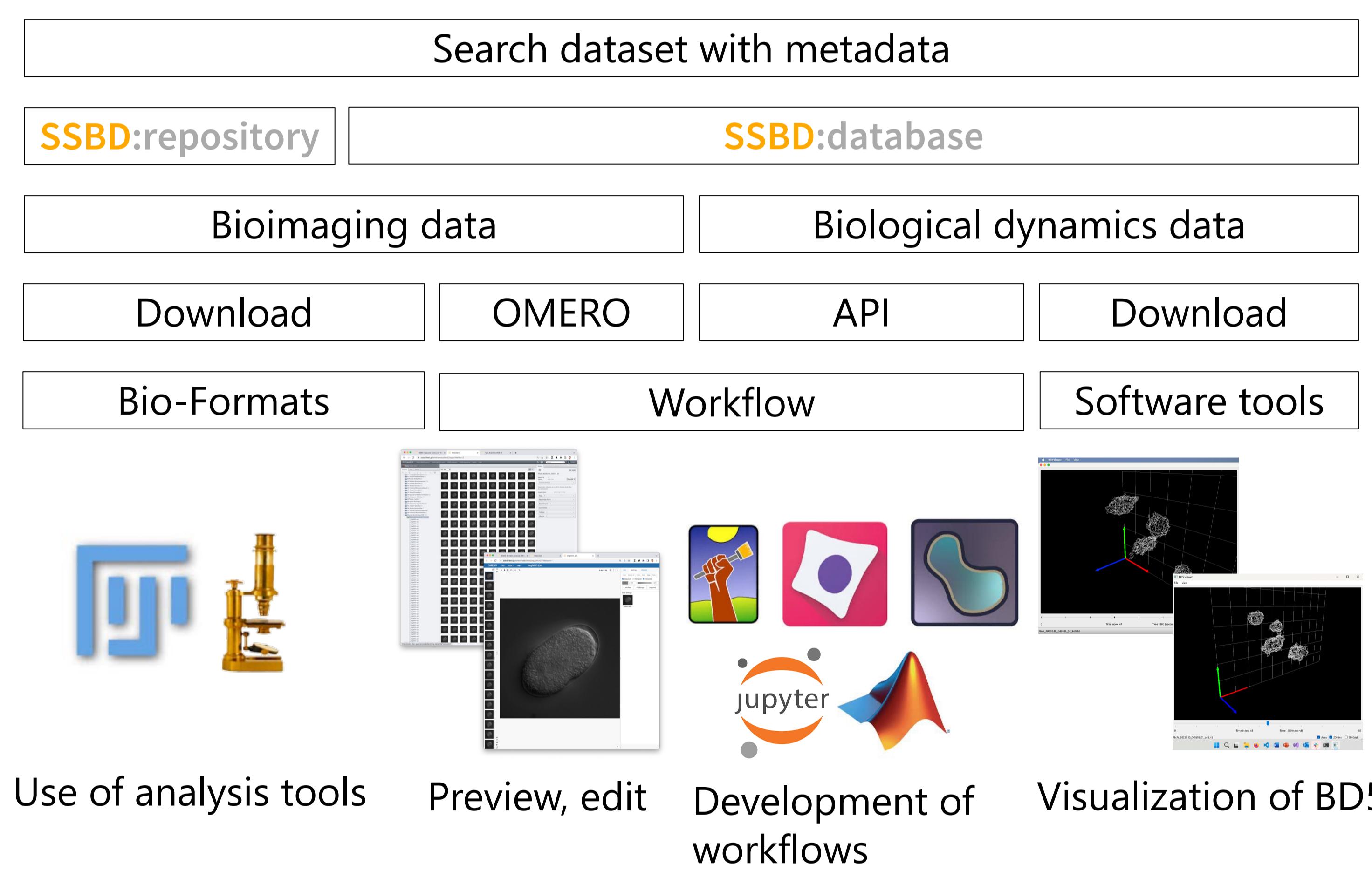
### Registration data to SSBD:repository

1. Request for data registration – e-mail to [ssbd-repos@ml.riken.jp](mailto:ssbd-repos@ml.riken.jp)
2. Filling in essential metadata in MS Excel format
3. Transfer of data and metadata via RIKEN Box, USB HDD, etc.
4. Cite issued DOI (Digital Objects Identifier) for the dataset in your paper



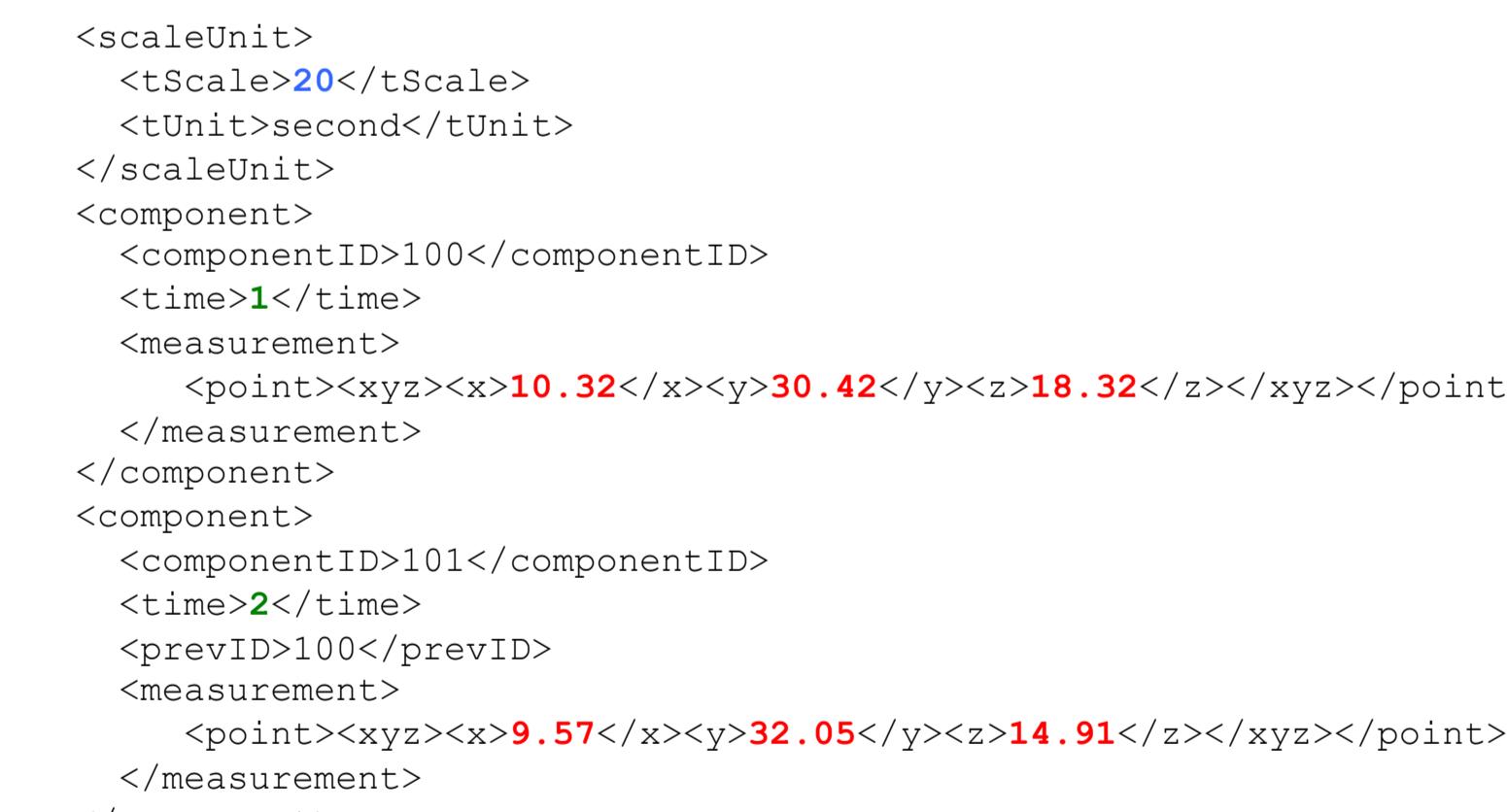
### Reuse of data

- Search dataset with metadata – biosample info, paper info, ontologies
- Access data and metadata via API
- Preview and process of bioimages on web via OMERO
- Visualization of biological dynamics data on web



### Data formats

- Bioimaging data – Original, unprocessed data (supported by Bio-Formats)
- Biological dynamics data – BDML/BD5 format



(Kyoda et al. 2015; 2020)

### Future works

- Easy data registration for SSBD:repository
- Integration to cloud systems
  - Support for OME-NGFF (Next Generation File Format)
  - Development of BD-zarr, next generation biological dynamics data format
- Sharing workflows, Provide for workflow engine

### References

- Tohsato, Y., et al. (2016). SSBD: a database of quantitative data of spatiotemporal dynamics of biological phenomena. *Bioinformatics* 32, 3471–3479. [10.1093/bioinformatics/btw417](https://doi.org/10.1093/bioinformatics/btw417).
- Kyoda, K., et al. (2015). Biological Dynamics Markup Language (BDML): an open format for representing quantitative biological dynamics data. *Bioinformatics* 31, 1044–1052. [10.1093/bioinformatics/btu767](https://doi.org/10.1093/bioinformatics/btu767).
- Kyoda, K., et al. (2020). BD5: An open HDF5-based data format to represent quantitative biological dynamics data. *PLoS One* 15, e0237468. [10.1371/journal.pone.0237468](https://doi.org/10.1371/journal.pone.0237468).
- Swedlow, J.R., et al. (2021). A global view of standards for open image data formats and repositories. *Nat Methods* 18, 1440–1446. [10.1038/s41592-021-01113-7](https://doi.org/10.1038/s41592-021-01113-7).
- Sarkans, U., et al. (2021). REMI: Recommended Metadata for Biological Images-enabling reuse of microscopy data in biology. *Nat Methods* 18, 1418–1422. [10.1038/s41592-021-01166-8](https://doi.org/10.1038/s41592-021-01166-8).
- Nelson, G., et al. (2021). QUAREP-LiM: A community-driven initiative to establish guidelines for quality assessment and reproducibility for instruments and images in light microscopy. *J Microsc* 284, 56–73. [10.1111/jmi.13041](https://doi.org/10.1111/jmi.13041).

