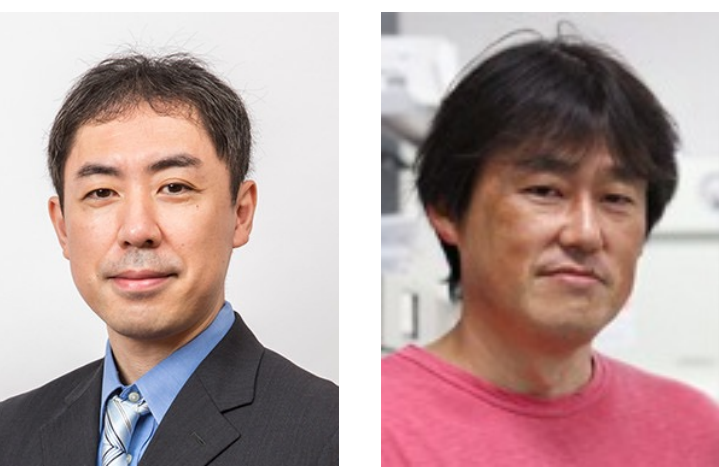


01

SSBD: A Global Bioimaging Data-Sharing Platform

SSBD: グローバルなバイオイメーシングデータ共有基盤

Hiroya Itoga¹, Koji Kyoda¹, Yuki Yamagata^{1,2}, Emi Fujisawa¹, Yuzheng Wu¹, Haruna Yamamoto¹, Tomohiro Nakamura¹, Chentao Wen¹, Ko Sugawara¹, Shuichi Onami^{1,3}

- ¹ RIKEN Center for Biosystems Dynamics Research, Japan
² RIKEN BioResource Research Center, Japan
³ RIKEN Information R&D and Strategy Headquarters, Japan

Check and Follow
“Bioimagram” on
Instagram!



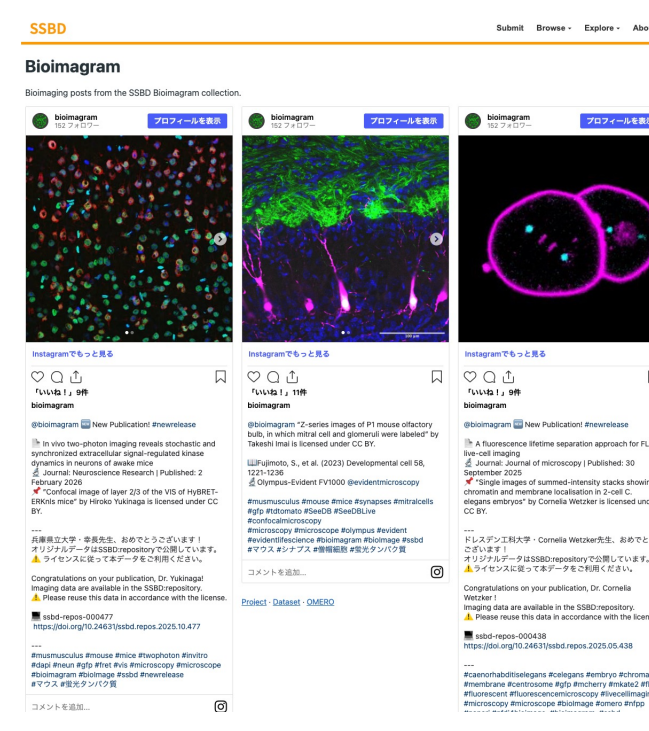
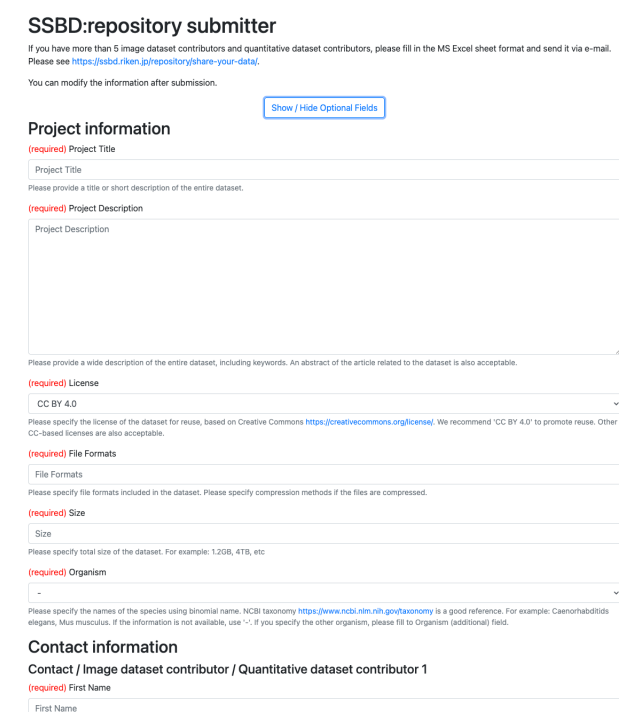
What is SSBD?

- SSBD^{1,2} is a publicly funded bioimaging data platform hosted by RIKEN, Japan.
- SSBD:repository supports rapid publication of diverse bioimaging data with essential metadata and DOIs.
- SSBD:database provides curated, richly annotated bioimaging and biological dynamics resources for broader and more advanced reuse.
- Together, these complementary services expand reusable bioimaging resources for human researchers and machines/AI systems.

Building reusable bioimaging resources since 2013

Data collection

- Easy submission – web submitter and communicator support
- Paper collection – 11,800+ papers; curator screening for bioimaging data
- Early bioimaging data discovery through bioRxiv
- Outreach via social media
- Luncheon seminars
- Tutorial videos



Data enrichment

- Selection and curation of highly reusable bioimaging data from SSBD:repository and collected papers, with rich annotation for interoperability and reusability
- Ontology-based, human-centric annotation with SSBD metadata v3 schema
- Machine-assisted dataset organization, not annotation

Data transfer

- Flexible large data transfer – RIKEN Box, R2DMS Nextcloud, Seafile CE
- Cloud-optimized image data format (OME-Zarr)
- Machine-accessible delivery and high-speed parallel download

Making SSBD resources discoverable and reusable worldwide

Discover globally

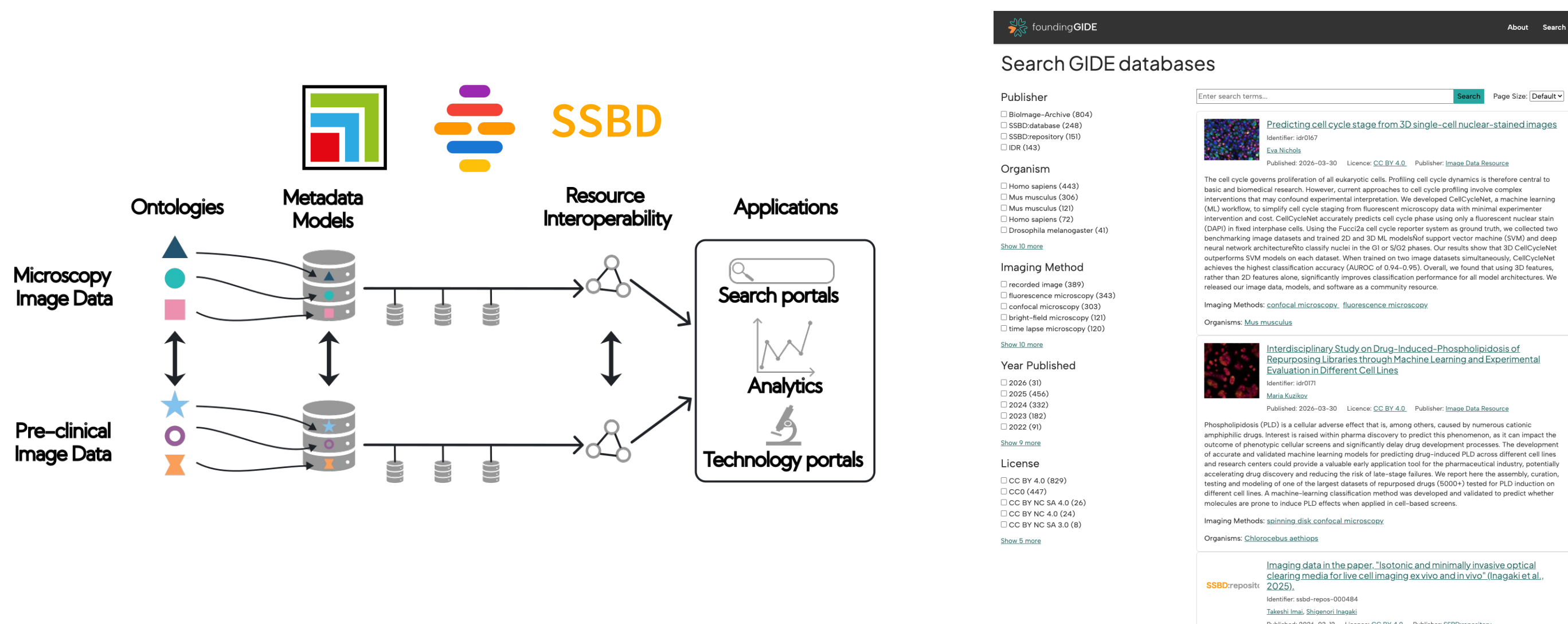
Global findability with GIDE Search

Outputs of the foundingGIDE⁵ project (2024/1 – 2026/6)

- GIDE Indexing and Search API
 - RO-Crate snapshots of metadata from SSBD, Image Data Resource (IDR), and BioImage Archive (BIA).
- GIDE BioImaging Search Portal
 - Search system to query projects/studies across SSBD, IDR, and BIA through a single interface.
 - Data remain hosted and served by their original repositories.

Beyond foundingGIDE

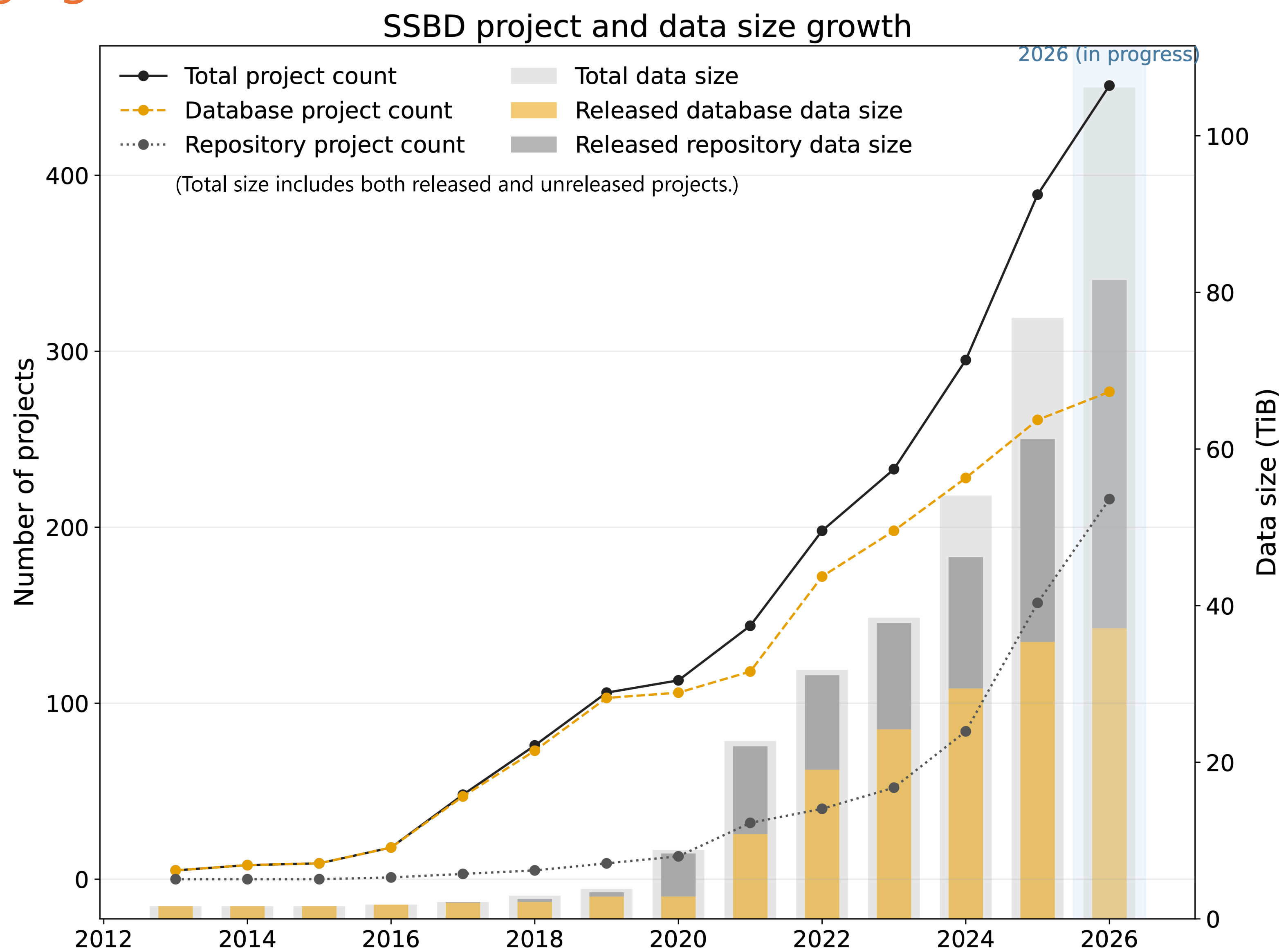
- Toward a worldwide bioimaging discovery initiative



Analyze globally – Future

Analysis environment powered by SSBD APIs

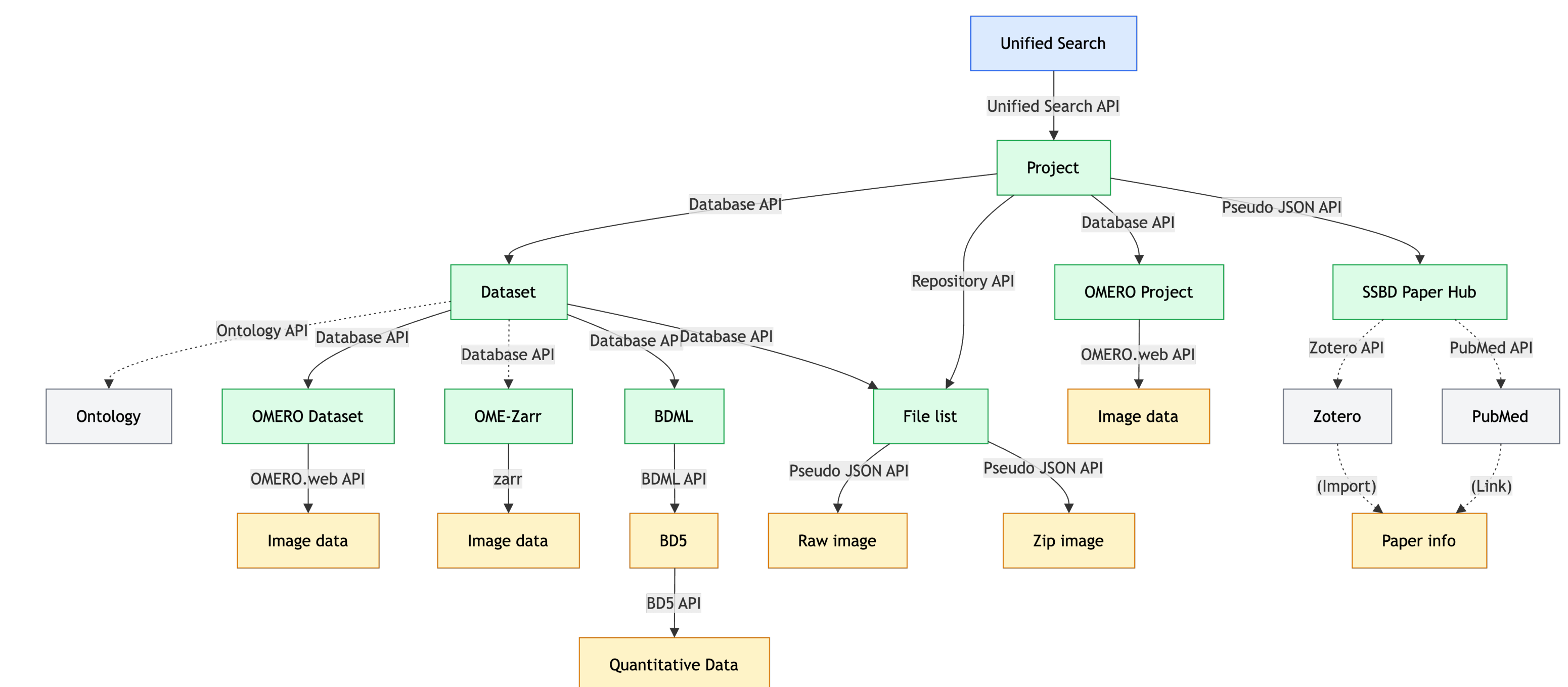
- Collaboration and integration with analysis tools
 - ELEPHANT: Tracking cell lineages in 3D by incremental deep learning⁷
 - 3DeeCellTracker⁸
- Provide a workflow execution environment for bioimaging data in SSBD
 - JupyterHub (Jupyter Notebook)
 - BIOIMERO – Bioimage Analysis in OMERO



Reuse globally

Local findability and global accessibility with Traversable SSBD

- Metadata is machine-searchable via Unified Search API across SSBD:database, SSBD:repository, and BDML³ metadata
- Metadata is machine-accessible via database, repository, and BDML APIs
- Data is machine-accessible via OME-Zarr⁶ schema, BD5⁴ API, and Pseudo-JSON file download API
- API responses link to the "next" resource, related SSBD resources, external databases, and ontology terms
- Traversal is enabled by OpenAPI specifications, machine-readable schemas, and links to related API endpoints
- A trial MCP server powered by the Traversable SSBD APIs is also available.



References

- Kyoda, K., Itoga, H., Yamagata, Y., Fujisawa, E., Wang, F., Miranda-Miranda, M., Yamamoto, H., Nakano, Y., Tohsato, Y., and Onami, S. (2025). SSBD: an ecosystem for enhanced sharing and reuse of bioimaging data. *Nucleic Acids Res* 53, D1716–D1723. doi:10.1093/nar/gkaf860, <https://ssbd.riken.jp/>.
- Tohsato, Y., Ho, K.H.L., Kyoda, K., and Onami, S. (2016). SSBD: a database of quantitative data of spatiotemporal dynamics of biological phenomena. *Bioinformatics* 32, 3471–3479. doi:10.1093/bioinformatics/btw417, <https://ssbd.qbic.riken.jp/>.
- Kyoda, K., Tohsato, Y., Ho, K.H.L., and Onami, S. (2015). Biological Dynamics Markup Language (BDML): an open format for representing quantitative biological dynamics data. *Bioinformatics* 31, 1044–1052. doi:10.1093/bioinformatics/btu767, <https://ssbd.qbic.riken.jp/bdml/>.
- Kyoda, K., Ho, K.H.L., Tohsato, Y., Itoga, H., and Onami, S. (2020). BD5: An open HDF5-based data format to represent quantitative biological dynamics data. *PLoS One* 15, e0237468. doi:10.1371/journal.pone.0237468.
- Onami, S., Kyoda, K., Yamagata, Y., and Itoga, H. (2025). foundingGIDE Deliverable D6.1: Report in metadata model overlap and gaps. doi:10.5281/zenodo.16794787.
- Moore, J., Basurto-Lozada, D., Besson, S., Bogovic, J., Bragantini, J., Brown, E.M., Burel, J.-M., Casas Moreno, X., de Medeiros, G., Diel, E.E., et al. (2023). OME-Zarr: a cloud-optimized bioimaging file format with international community support. *Histochem Cell Biol* 160, 223–251. doi:10.1007/s00418-023-02209-1.
- Sugawara, K., Çevrim, Ç., and Averof, M. (2022). Tracking cell lineages in 3D by incremental deep learning. *eLife* 11, e69380. doi:10.7554/eLife.69380.
- Wen, C., Miura, T., Voleti, V., Yamaguchi, K., Tsutsumi, M., Yamamoto, K., Otomo, K., Fujie, Y., Teramoto, T., Ishihara, T., et al. (2021). 3DeeCellTracker, a deep learning-based pipeline for segmenting and tracking cells in 3D time-lapse images. *eLife* 10, e59187. doi:10.7554/eLife.59187.