



Artificial Intelligence for Image Data Analysis in the Life Sciences

Beatriz Serrano-Solano

Scientific Project Manager at Euro-BioImaging ERIC

AI4Life WP7 lead (Communication, Outreach & Training)



AI4Life has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number 101057970.



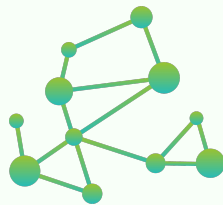
The European Research Infrastructure for
Biological & Biomedical Imaging

- ✓ **Distributed** Research Infrastructure with nodes across Europe
- ✓ **Nodes** are internationally recognised imaging facilities
- ✓ **Publicly funded**, not-for-profit
- ✓ Provides **open access** to biological and biomedical imaging technologies and data services



17

ERIC MEMBERS
(16 COUNTRIES & EMBL)



35

NODES



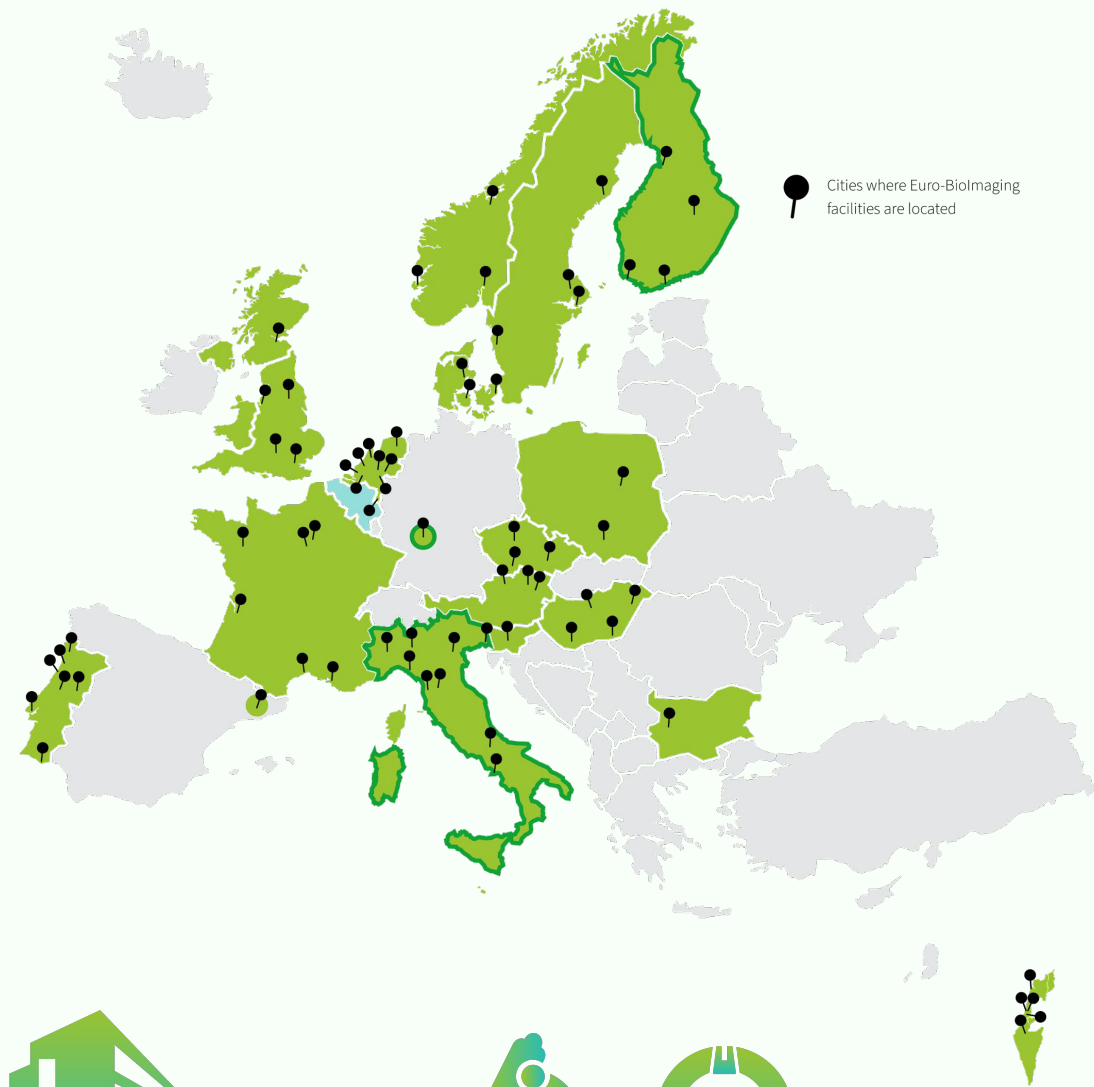
170+

SITES

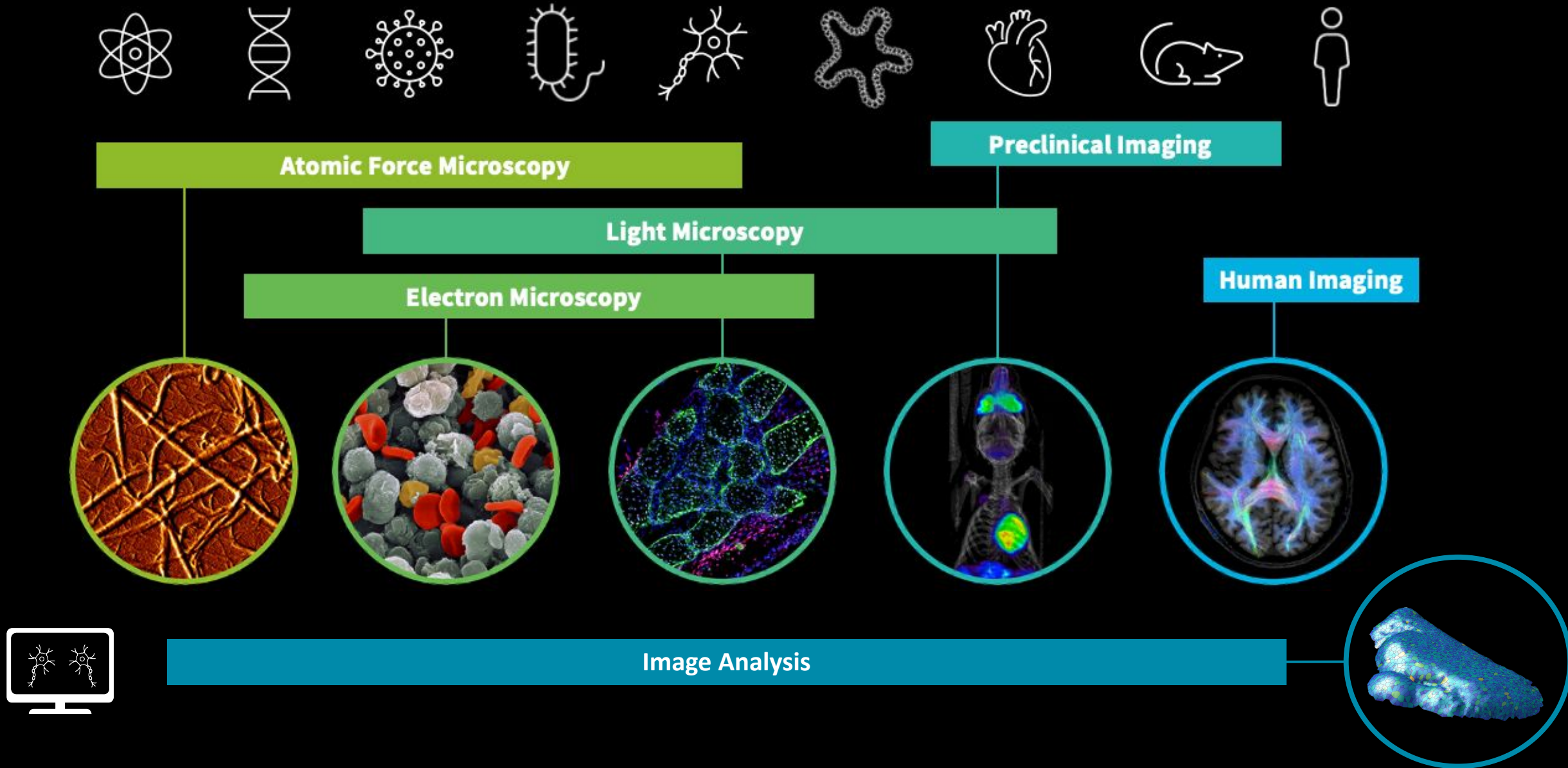


100+

TECHNOLOGIES



The Euro-Biolmaging technology portfolio

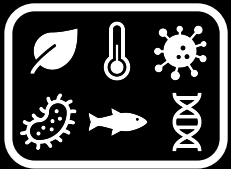


F_{indable} A_{ccessible} I_{nteroperable} R_{eusable} Challenges



Size

- Single experiment (Giga to Tera bytes)
- Multiple dimensions (2D to 4D)
- Longitudinal studies, live imaging



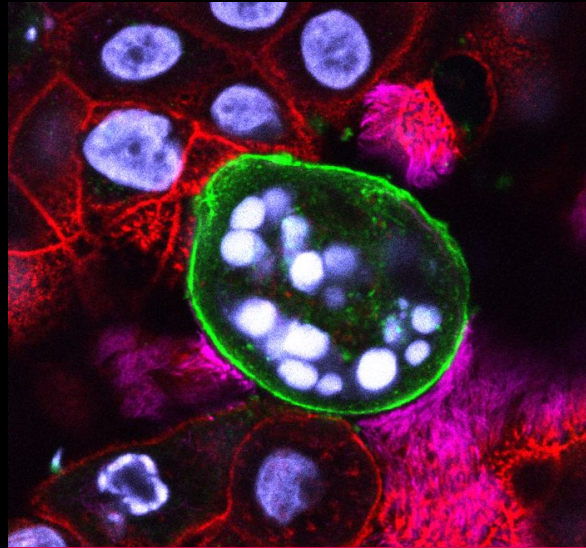
Complexity

- Multimodal and correlative datasets
- Various areas of application
- Multicontrast and multiparametric



Format

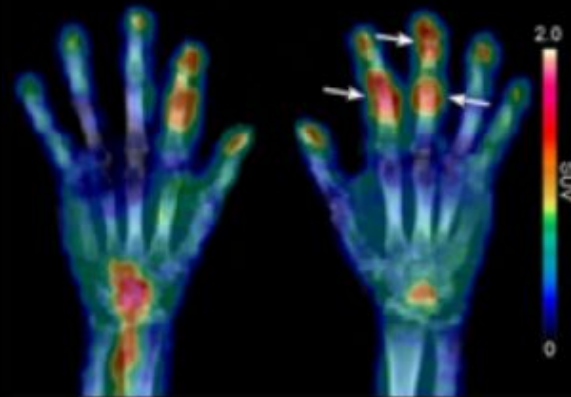
- Diversity in instrumentation, proprietary file formats
- For binary data and metadata



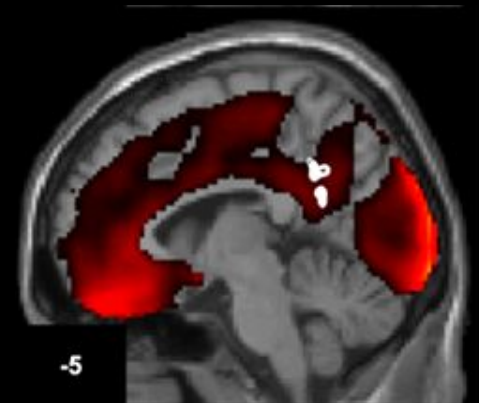
France BioImaging node



Portuguese Platform of BioImaging



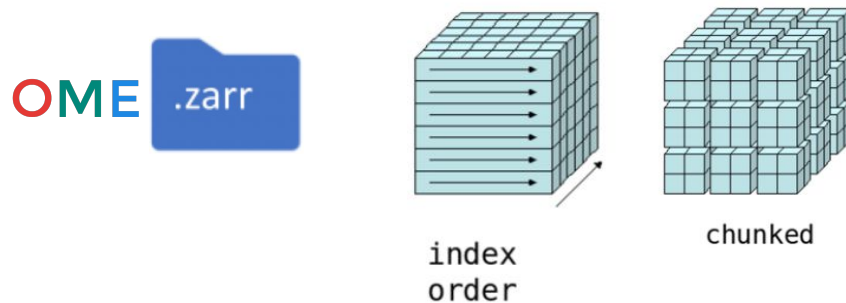
Finnish Biomedical Imaging node



Brain Imaging Network node

File formats: OME-Zarr

- **Standard** image format
 - Common layout for binary data + **metadata**
 - Higher access **speed**, especially for remotely stored datasets
- Euro-Biolmaging **services**:
 - Tool development to support OME-Zarr:
<https://github.com/orgs/Euro-Biolmaging/repositories>
 - Guidance for submission of data to public repositories in OME-Zarr format



nature > technology features > article

TECHNOLOGY FEATURE | 02 October 2023

How open-source software could finally get the world's microscopes speaking the same language

A plethora of standards mean shareable and verifiable microscopy data often get lost in translation. Biologists are working on a solution.

Michael Brooks

Home > Histochemistry and Cell Biology > Article

OME-Zarr: a cloud-optimized bioimaging file format with international community support

Original Paper | Open Access | Published: 10 July 2023 | 160, 223–251 (2023)

Download PDF | You have full access to this open access article

Brief Communication | Open Access | Published: 29 November 2021

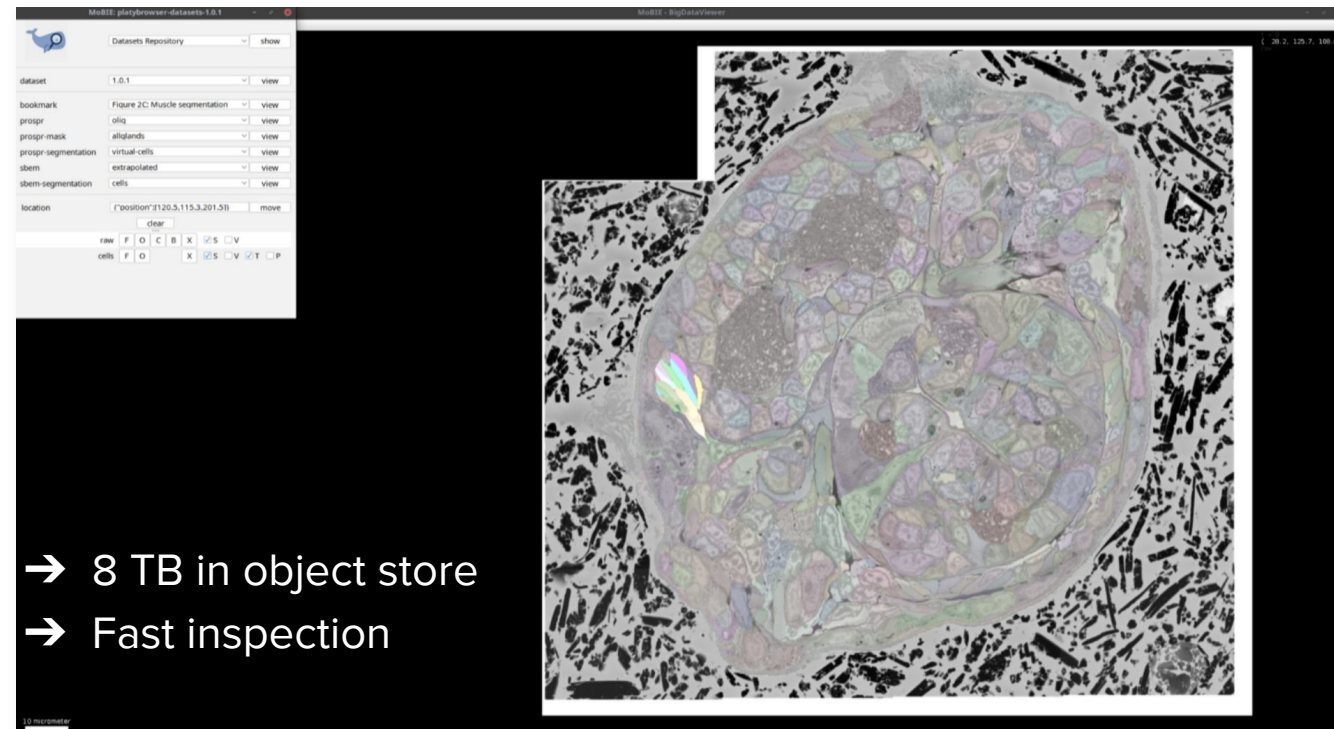
OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies

Josh Moore, Chris Allan, Sébastien Besson, Jean-Marie Burel, Erin Diel, David Gault, Kevin Kozłowski, Dominik Lindner, Melissa Linkert, Trevor Manz, Will Moore, Constantin Pape, Christian Tischer & Jason R Swedlow

Nature Methods 18, 1496–1498 (2021) | Cite this article

9267 Accesses | 16 Citations | 81 Altmetric | Metrics

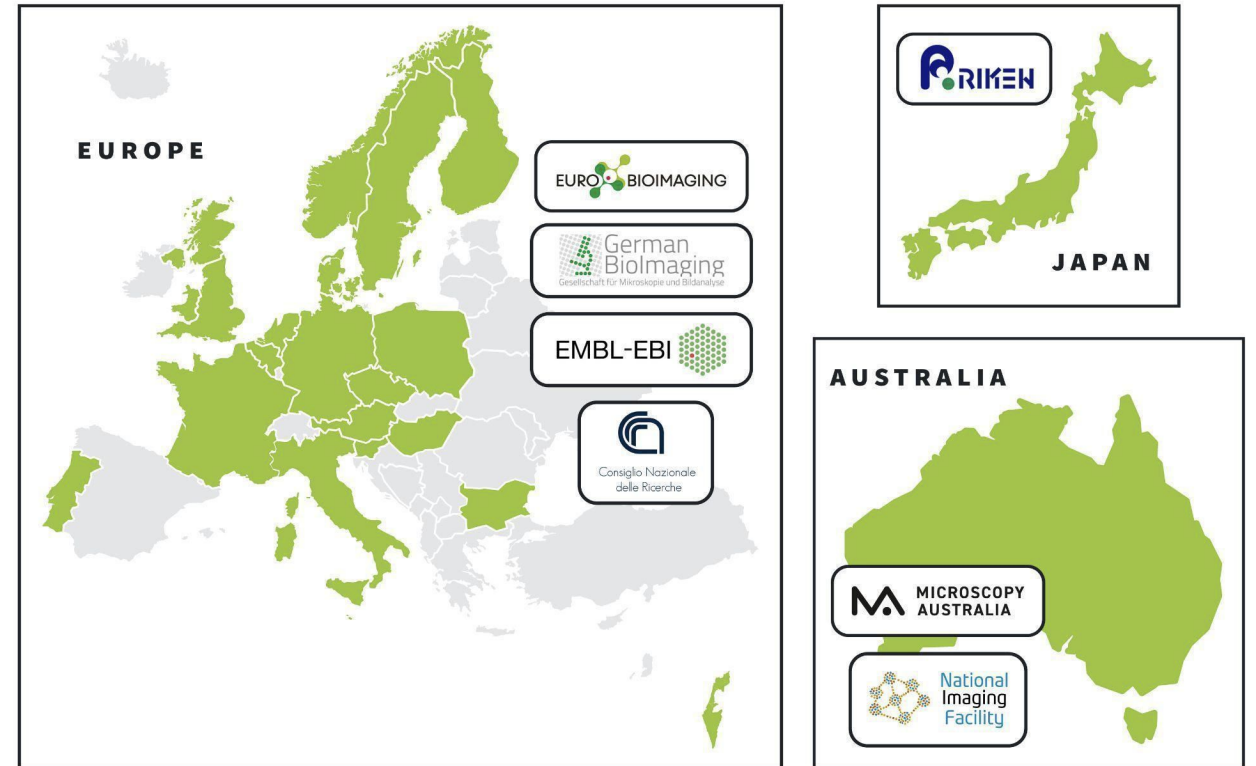
esson, John Bogovic, Jordão Bragantini, Eva M. Javo de Medeiros, Erin E. Diel, David Gault, Satrajit S. Bartley, Dave Horsfall, Mark S. Keller, Mark Iji Kyoda, Albane le Tournoux de la Villegeorges, Kert, Joel Lüthi, Jeremy Maltin-Shepard, Trevor Lange, Khaled Mohamed, William Moore, Nils Constantin Pape, Lucas Pelkmans, Tobias Pietzsch, meeuul Samee, Nicholas Schaub, Hythem Sidky, Christian Tischer, Daniel Toloudis, Isaac Virshup, ces Wong, Kevin A. Yamauchi, Omer Bayraktar, Beth Hotaling, Shuichi Onami, Loic A. Royer, Stephan dlow — Show fewer authors



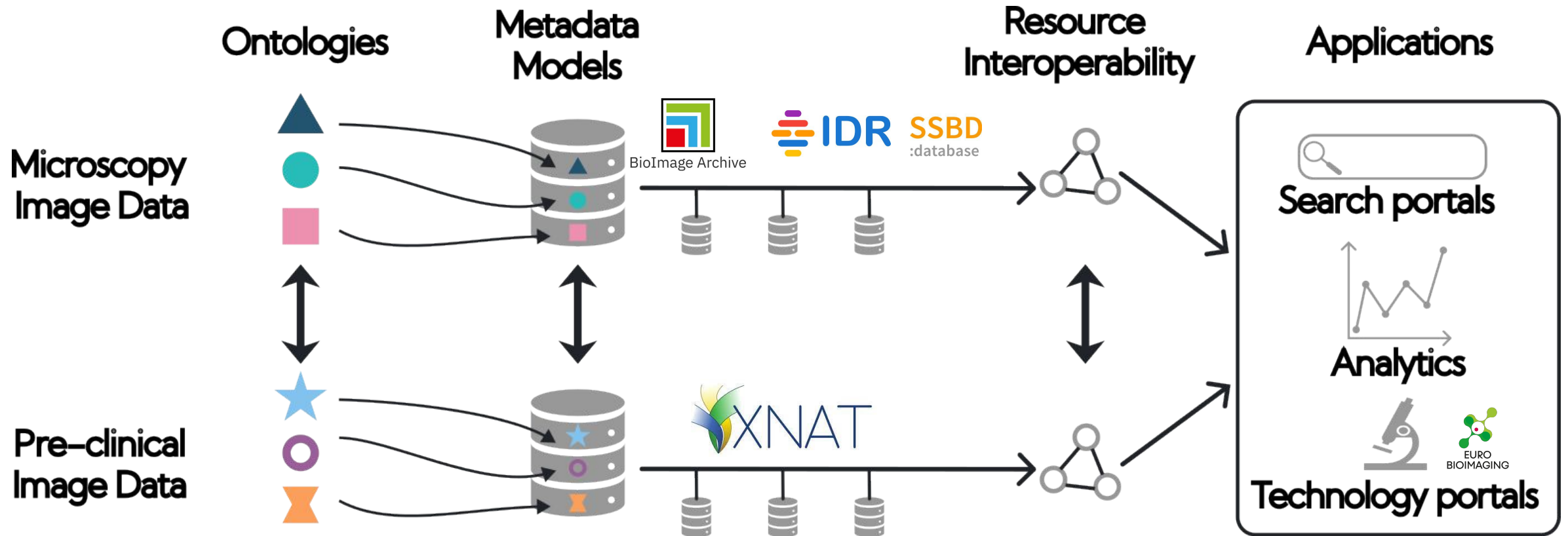
Founding**GIDE**: Founding a **Global Image Data Ecosystem**

Laying strong foundation of an ecosystem for **image data exchange** based on global coordination of technical developments among data infrastructures and communities

- Global coordination among diverse imaging resources & communities
- Concerted development of Ontologies and Metadata models
- Adoption of outputs by global image data resources
- Interoperable solutions for microscopy and pre-clinical data
- Community recommendations FAIR image data management



Founding **GIDE**: Founding a **Global Image Data Ecosystem**



AI4Life

a Quick Overview

AI4Life overview



10

PARTNERS

8

COUNTRIES

3

YEARS

4M

EUROS

Bridging the Life & Computer Science Communities



The goal is to radically **reduce the disparity** between the theoretical applicability and the practical use of AI-based image analysis methods in the life sciences. AI4Life will bridge the gap between the two rapidly developing fields of **AI/ML methods** development and **biological imaging**.

Most visible, user-facing infrastructure:

- Model Zoo (www.bioimage.io) – <https://www.biorxiv.org/content/10.1101/2022.06.07.495102v1>
- BioImage Archive
- Open Calls
- Challenges

AI4Life objectives

1

**Democratized
availability** of
AI-based image analysis
methods

2

Establish standards for the
submission, storage and
FAIR access

3

**Simple model deployment,
sharing, and dissemination**
through a new
developer-facing service

4

Organize **Open Calls and
Challenges** for image
analysis problems

5

Empower **common image
analysis** platforms with **AI
integration**

6

**Organizing outreach and
training** events i.e. image
analysis courses/workshops
and participation in
international conferences

AI4Life structure

WP1 – PROJECT MANAGEMENT

WP2 – USER SERVICES &
COMPUTING INFRASTRUCTURES

WP3 – DIRECT SUPPORT

WP4 – CONTRIBUTOR SERVICES

WP5 – DATA, MODEL & COMPUTING
STANDARDS

WP6 – SUPPORT FOR OPEN CALLS,
CHALLENGES & NEW SERVICES

WP7 – COMMUNICATION, OUTREACH
& TRAINING

AI4Life PARTNERS

Community Partners

 BiolImage.IO

 #ZeroCostDL4Mic

 deepImageJ



 ImJoy

 ilastik

 HPA

Project Partners

EMBL



EMBL-EBI



 EMPHASIS

 EURO BIOIMAGING

eu:openscreen

 HUMAN
TECHNOPOLE

 INSTITUTO
GULBENKIAN
DE CIÊNCIA

 instruct
ERIC



uc3m | Universidad
Carlos III
de Madrid



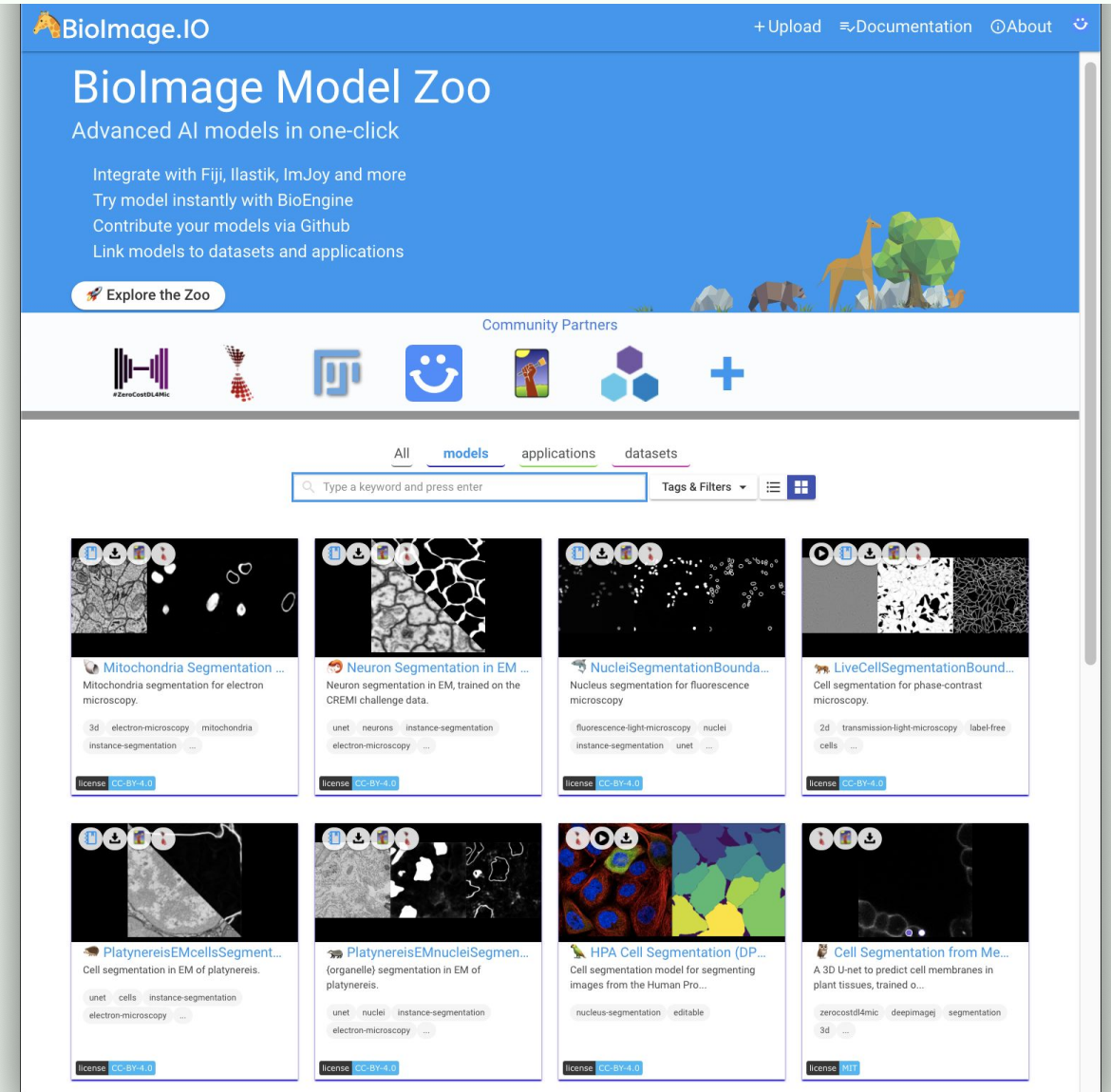
Advanced AI models in one click

Bioimage Model Zoo

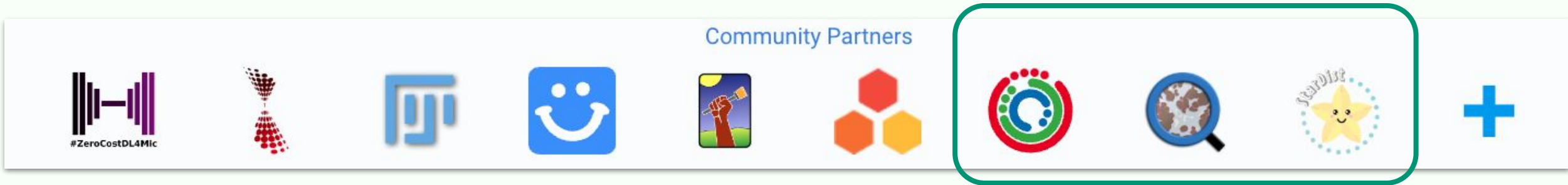
A repository for FAIR AI models

BiolImage Model Zoo

- **Community-driven**
- **Open source**
- Community-based process for **contributing models**
 - Cross-compatible models among the consumer software, should always run on at least one
 - Model should be well documented
 - Model should be public and reusable under the chosen licensing conditions
- Users choose the consumer software & download
- Make **sharing** and **application** of **pre-trained neural networks** available and easy!



Community Partners



Icy

icy.bioimageanalysis.org

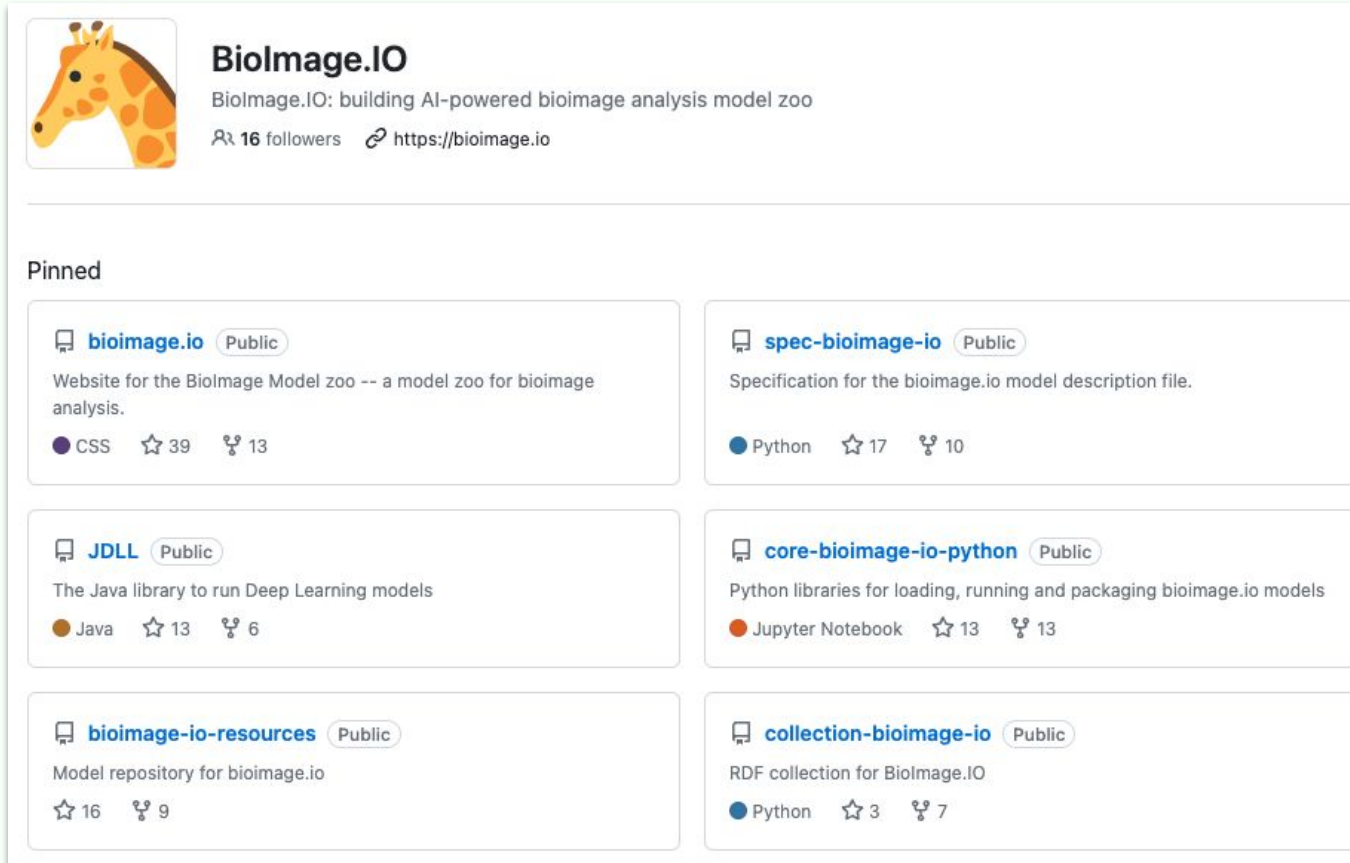
QuPath

qupath.github.io

StarDist

stardist.github.io

AI Model Specification



BioImage.IO
BioImage.IO: building AI-powered bioimage analysis model zoo
16 followers <https://bioimage.io>

Pinned

- bioimage.io** (Public)
Website for the BioImage Model zoo -- a model zoo for bioimage analysis.
CSS 39 13
- spec-bioimage-io** (Public)
Specification for the bioimage.io model description file.
Python 17 10
- JDLL** (Public)
The Java library to run Deep Learning models
Java 13 6
- core-bioimage-io-python** (Public)
Python libraries for loading, running and packaging bioimage.io models
Jupyter Notebook 13 13
- bioimage-io-resources** (Public)
Model repository for bioimage.io
16 9
- collection-bioimage-io** (Public)
RDF collection for BioImage.IO
Python 3 7

A model contains

- Trained model
- Example input image
- Example output
- Example cover image for visual representation in the GUI
- RDF (Resource Description File) specification
 - YAML file
 - Mandatory and optional fields
 - Architecture
 - input/output format
 - Preprocessing steps
 - Performance metrics



RDF Models
(collection.json)

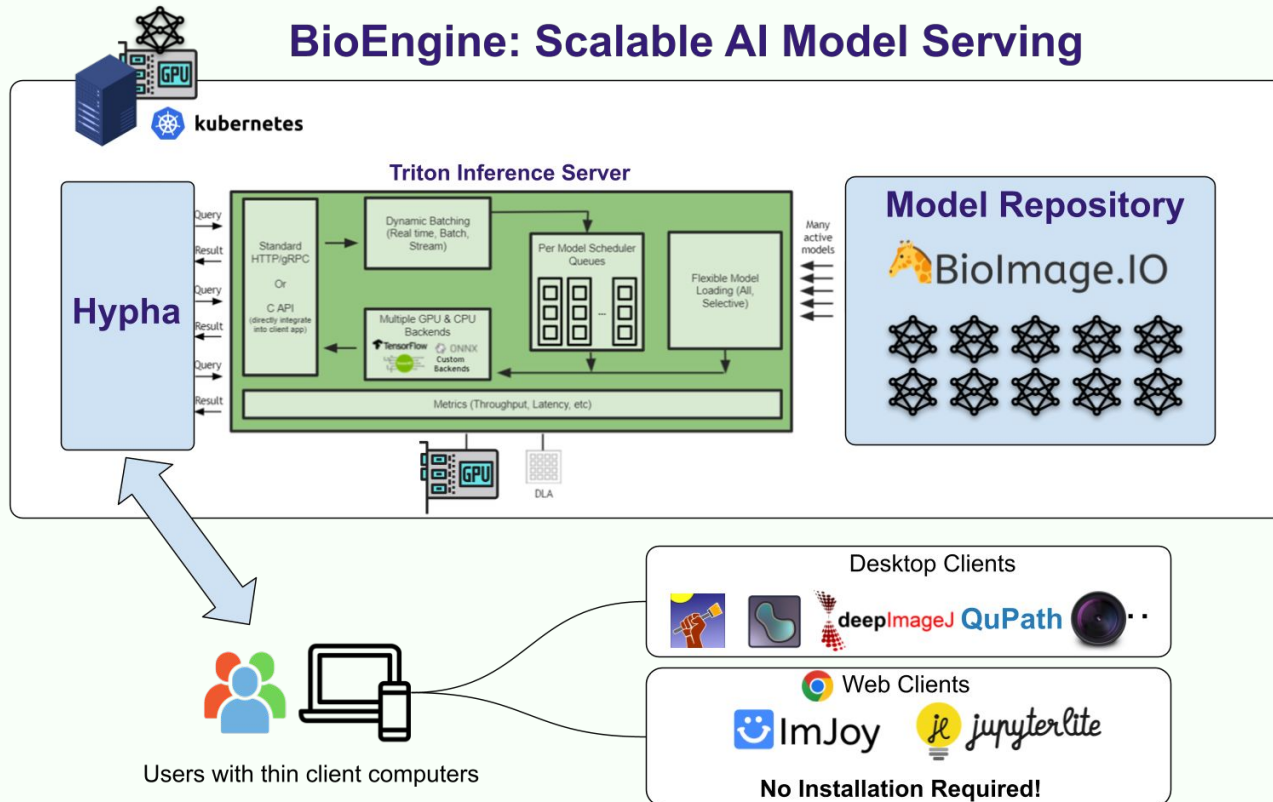


Validator scripts
(generate, etc.)

<https://github.com/bioimage-io>



BioEngine Model Runner



- Run models in the browser with your own images (training + inference)
- Shared GPU resources
- Local or cloud deployment
- Multi-user
- Multiple tools

<https://github.com/bioimage-io/bioengine>



Training & Support

- Documentation










<https://bioimage.io/docs/>

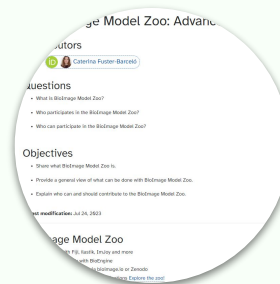
- Training materials available at the Galaxy Training Network

<https://training.galaxyproject.org/training-material/topics/ai4life/>

- Help Desk

<https://ai4life.eurobioimaging.eu/help/>

Lesson	Slides
Biolmage Model Zoo: Advanced models in one-click bioimageio image analysis	  
Contribute to the Biolmage.IO models bioimageio image analysis	  
Welcome to the Biolmage Model Zoo bioimageio image analysis	  



Plain text

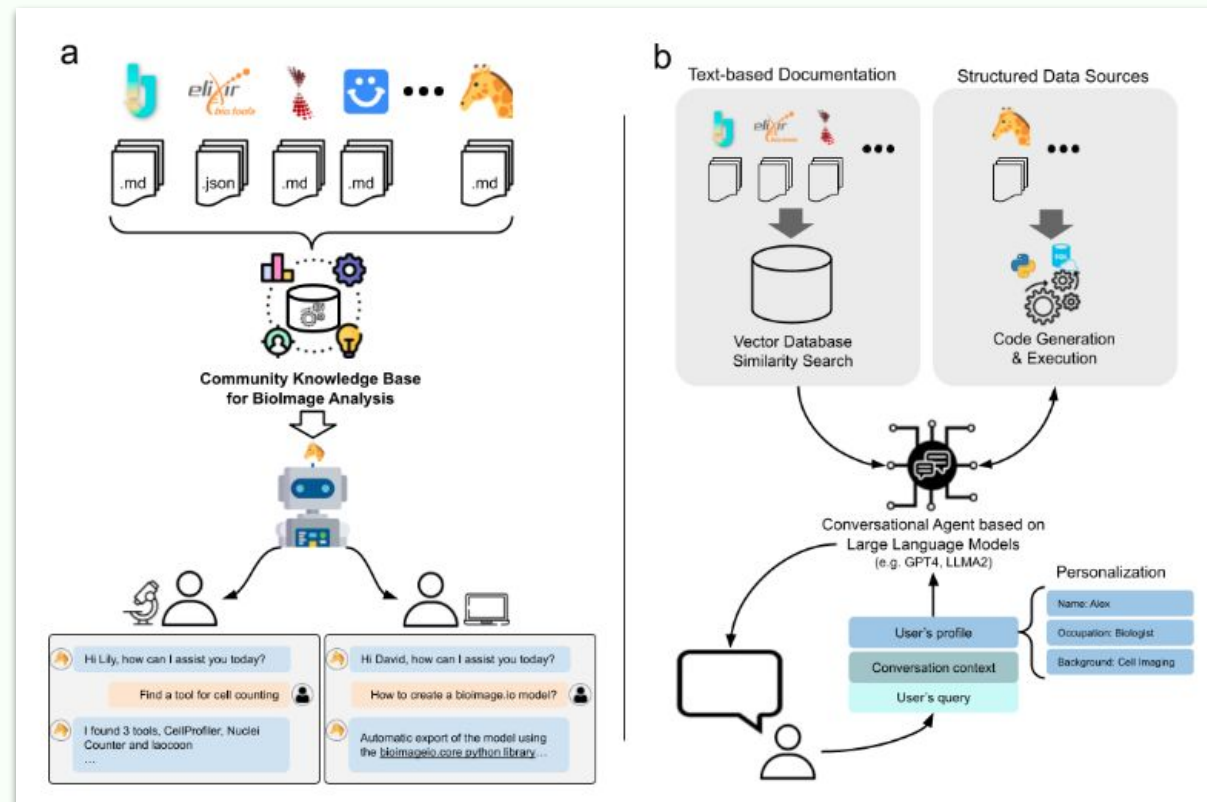


Slides



Videos

Biolmage.io Chatbot



Preprint now available!

23/10/2023

Biolmage.IO Chatbot: A Personalized Assistant for Biolmage Analysis Augmented by Community Knowledge Base

Wanlu Lei^{1,2}, Caterina Fuster-Barceló^{3,4}, Arrate Muñoz-Barrutia^{3,4}, Wei Ouyang^{1,2}

¹Department of Intelligent Systems, KTH Royal Institute of Technology, Stockholm, Sweden

²Ericsson Inc., Santa Clara, CA, USA

³Bioengineering Department, Universidad Carlos III de Madrid, Leganes, Spain

⁴Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

⁵Department of Applied Physics, Science for Life Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden

*Correspondence: weio@kth.se

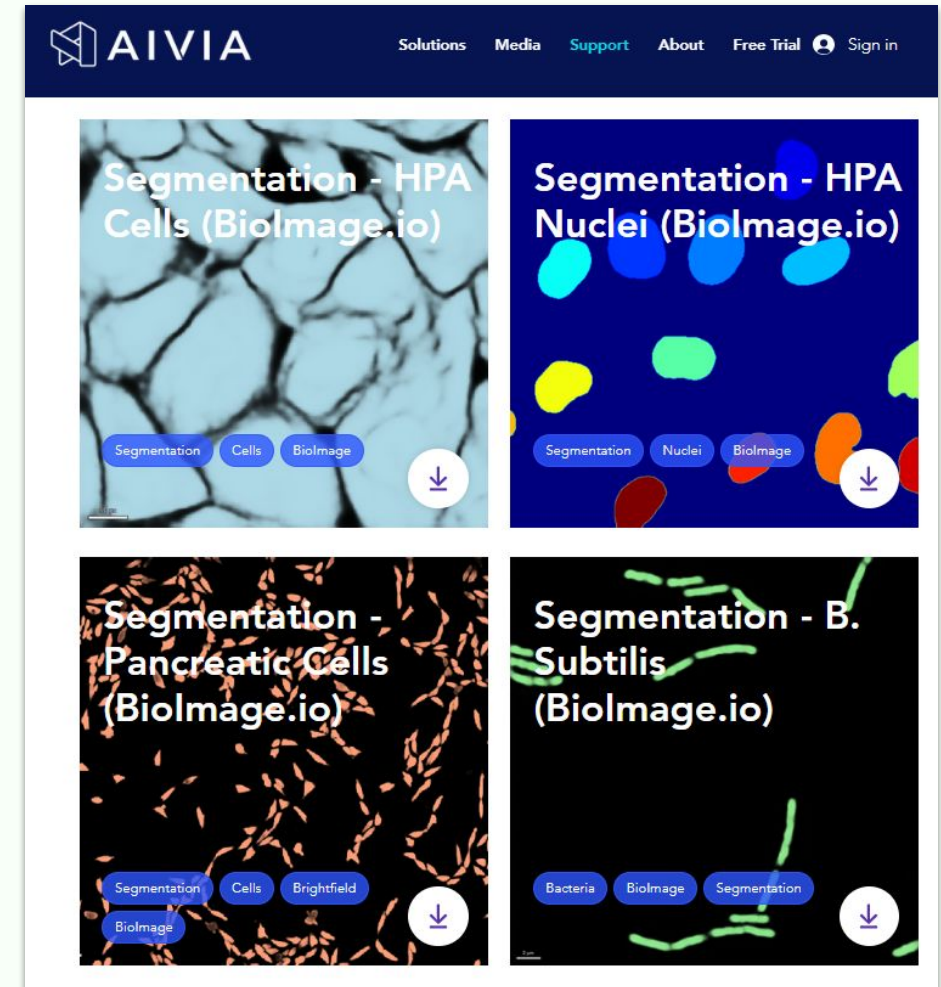
- Trained with the documentation of the BMZ and community partners

<https://github.com/bioimage-io/bioimageio-chatbot>

Collaboration with Industry

- Leica Microsystems software
- Model conversion for interoperability with AIVIA
- Ongoing conversations with more industry partners

<https://ai4life.eurobioimaging.eu/ai4life-leica-collaboration/>

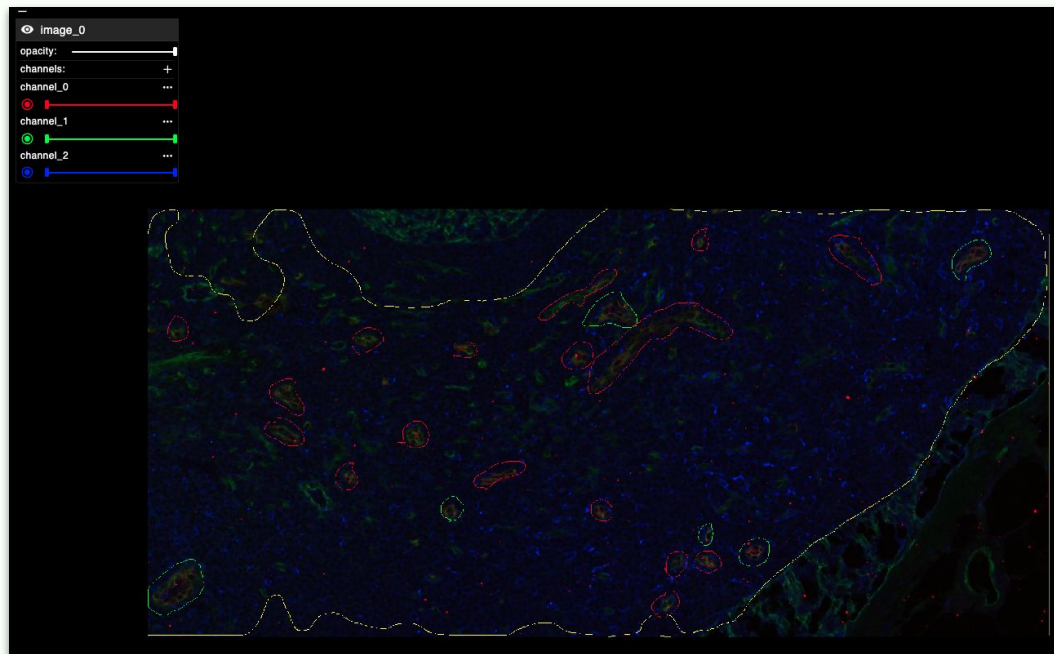




Bioimage Archive


A repository for image data

AI datasets at the BIA



Fluorescence images and geometrical annotations and classification of high endothelial venules (HEVs) in tumor-draining lymph nodes

(Visualisation using vizarr)

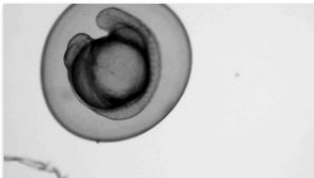
 **BioImage Archive**

Home | Browse | Submit | **Galleries** | Help | REMBI Help | Policies | About us

ALPHA **BioImage Archive AI datasets**
A selection of AI related studies

This is a collection of AI/Machine Learning datasets from the BioImage Archive from which one or more images have been converted to OME-Zarr. It is intended to present the AI related datasets with relevant tags and visualisation of images from the archive's collection, and to provide easy access to AI datasets and encourage tool development.

[Check here for how to contribute](#)

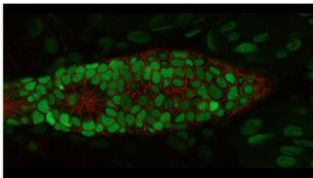


S-BIAD531 →

Live widefield images of zebrafish embryo with segmentation masks and class labels for temporal development

1142 images , 1062 annotations

Tags: segmentation masks, class labels, time series

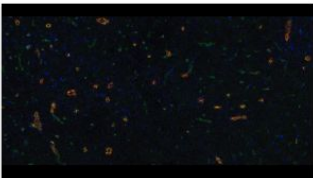


S-BIAD599 →

3D zebrafish embryo images with single-cell segmentation and point cloud-based morphometry

584 images , 196 annotations

Tags: segmentation masks, 3D

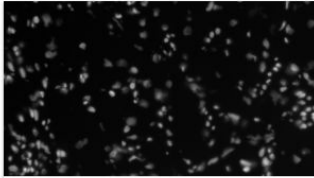


S-BIAD463 →

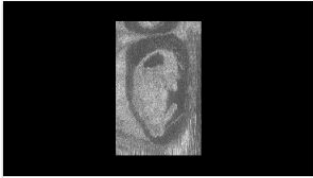
Fluorescence images and geometrical annotations and classification of high endothelial venules (HEVs) in tumor-draining lymph nodes

1285 images , 729 annotations

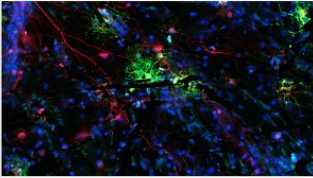
Tags: geometrical annotations, weak annotations, bounding boxes, class labels



S-BIAD634 →



S-BIAD686 →



S-BIAD493 →

<https://www.ebi.ac.uk/bioimage-archive/galleries/AI.html>

Towards FAIR AI image data

Virtual workshop on image annotations



- Image **annotations** are critical for **reproducibility**, and also frequently very time intensive to produce
- **Sharing/reuse** of **annotations** is hugely valuable, e.g for methods development
- What can we do to make these **annotations FAIR** and open?

The collage displays four entries from the BioStudies/BIOIMAGES database:

- Entry 1 (S-BIAD161):** "Machine learning reveals mesenchymal breast carcinoma cell adaptation in response to matrix stiffness". Release Date: 2 July 2021. Modified: 1 July 2021. Authors: Vlada S. Rozova, Ayad G. Anwer, Anna E. Guller, Hamidreza Aboulkheyr Es, Zahra Khabir, Anastasiya I. Sokolova.
- Entry 2 (S-BSST448):** "Citizen science, cells and CNNs – deep learning for automatic segmentation of the nuclear envelope in electron microscopy data, trained with volunteer segmentations". Release Date: 28 July 2020. Modified: 11 August 2020. Authors: Helen Spiers, Harry Songhurst, Luke Nightingale, Joost de Folter, Roger Hutchings, Christopher J Peddie, Anne Weston.
- Entry 3 (S-BSST641):** "In support of: Inter-laboratory automation of the in vitro micronucleus assay using imaging flow cytometry and deep learning". Release Date: 3 May 2021. Modified: 3 May 2021. Authors: Dr John W. Wills, Professor Paul Rees, Dr George Johnson.
- Entry 4 (S-BIAD463):** "Automated detection of vascular remodeling in tumor-draining lymph nodes by the deep learning tool HEV-finder". Release Date: 20 May 2022. Modified: 20 May 2022. Authors: Tove Bekkhus, Christophe Avenel, Anna Klemm, Carolina Wählby, Maria Ulvmar.

The bottom screenshot (S-BIAD463) also includes a "Data files" table:

Name	Size	Section	Description
1.png	2.6 MB	Study Component	RGB image (see my_submission/Tr for list of annotatic RGB image (see

Metadata

Study

- Description
- License
- Versioning
 - Current
 - Previous
- Usage: AI models using the dataset (BMZ)

Source image

WHAT

- Semantic labels (e.g. "cell wall")
 - Ontologies or controlled vocabularies
- REMBI (http://bit.ly/rembi_v1)



Provenance

- Authors (original annotators) **WHO**
- Timestamps **WHEN**
- Method **HOW**
 - Software
 - Protocols
 - QA
 - Experts or crowdsourced?
 - Produced by humans or software?
- Confidence level (e.g. #years of experience)
- Predicted by an algorithm or curated from a prediction

Annotations

WHERE

- Link to the original data
- Type: counts, which images
- Spatial information
 - Non-pixel annotations (e.g. counts of items in a ROI)
 - Geometrical primitives (points, lines, contours, polygons)
 - Segmentations (label maps, area/volume)
- Transformations
 - Rotations
 - Translations
- Linking labels across time points (for tracking)

Compiled by Teresa Zulueta-Coarasa

FAIR AI: Recommendations

Reduce number of formats

NGFF/OME-Zarr

COCO

GeoJSON

EMDB-SFF

CSV/TSV



API

Data browsing

Link to community tools

Metadata search

Dataset curation

Reach out to journals



Make data accessible

Standardise metadata

Annotators

Annotation type

Annotation method

Annotation coverage

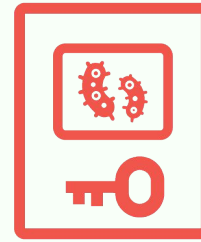
Annotation criteria

Association with source image

Confidence level

Transformations

Spatial information



Preserve annotator credit

Organise annotation events

Acknowledge exceptional contributors



Encourage data production

<https://ai4life.eurobioimaging.eu/summary-fair-ai-workshop/>

Opportunities

Workshops & Hackathons

[https://ai4life.eurobioimaging.eu/**events**/](https://ai4life.eurobioimaging.eu/events/)

Open Calls

[https://ai4life.eurobioimaging.eu/**open-calls**/](https://ai4life.eurobioimaging.eu/open-calls/)

Challenges

[https://ai4life.eurobioimaging.eu/**challenges**/](https://ai4life.eurobioimaging.eu/challenges/)

Workshops & Hackathons

<https://ai4life.eurobioimaging.eu/events/>
<https://ai4life.eurobioimaging.eu/news/>



First AI4Life Hackathon

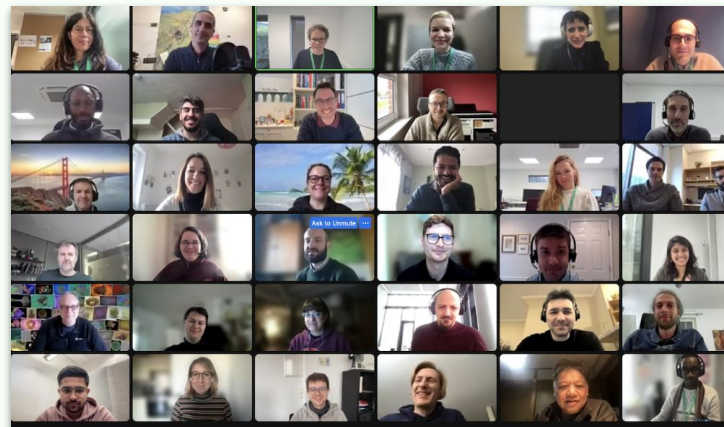


Hackathon Deep Learning in Java



Hackathon Web & Cloud infrastructure for AI-powered bioimage analysis

Workshop
FAIR AI



AI4Life & BioImage Archive FAIR AI Workshop - Summary of recommendations

Workshop participants

Summary of recommendations gathered at the workshop "AI4Life & BioImage Archive FAIR AI".



First AI4Life Open Call

OPEN CALL for Bioimage Analysis Support



- ☐ I need help analysing image data
- ☐ I need help creating training data
- ☐ I need help applying Deep Learning to my research data

Apply by
March 31



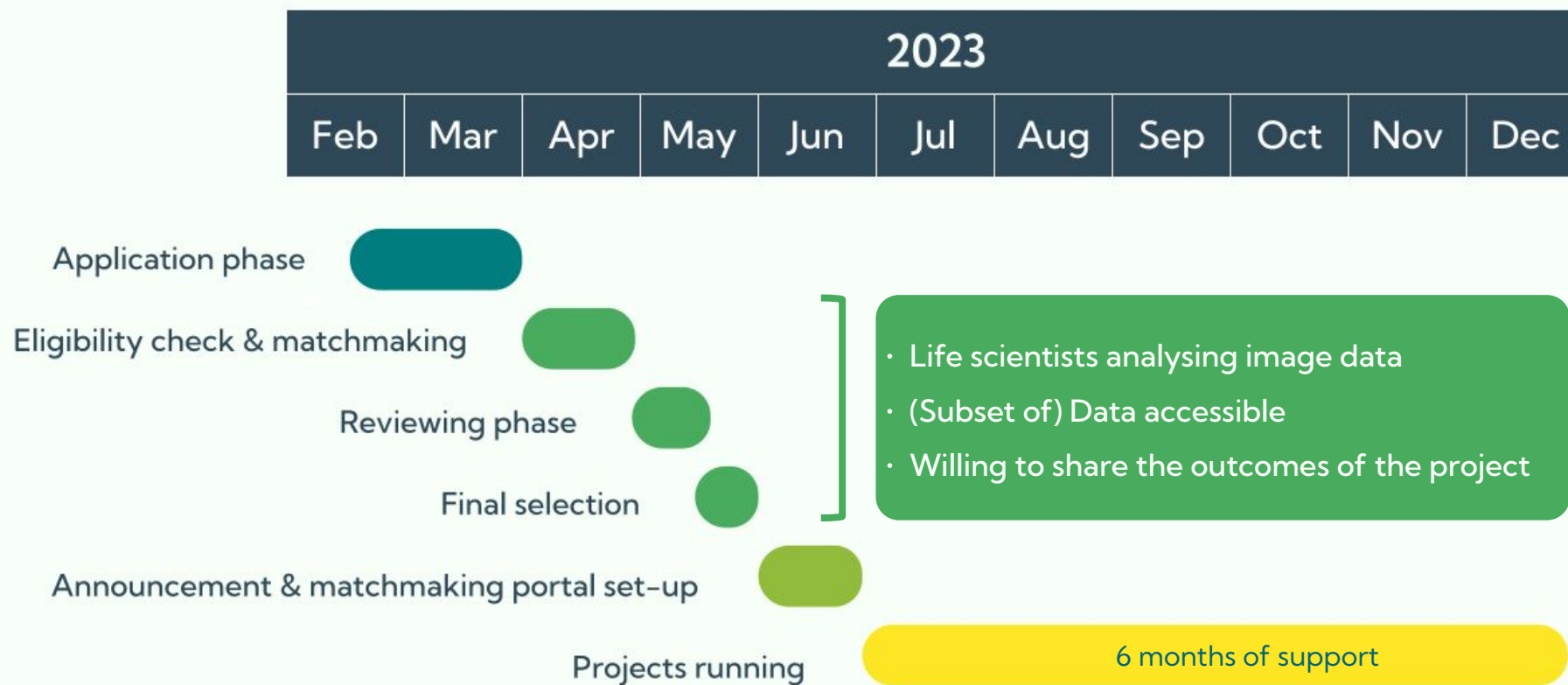
AI4Life



ai4life@eurobioimaging.eu

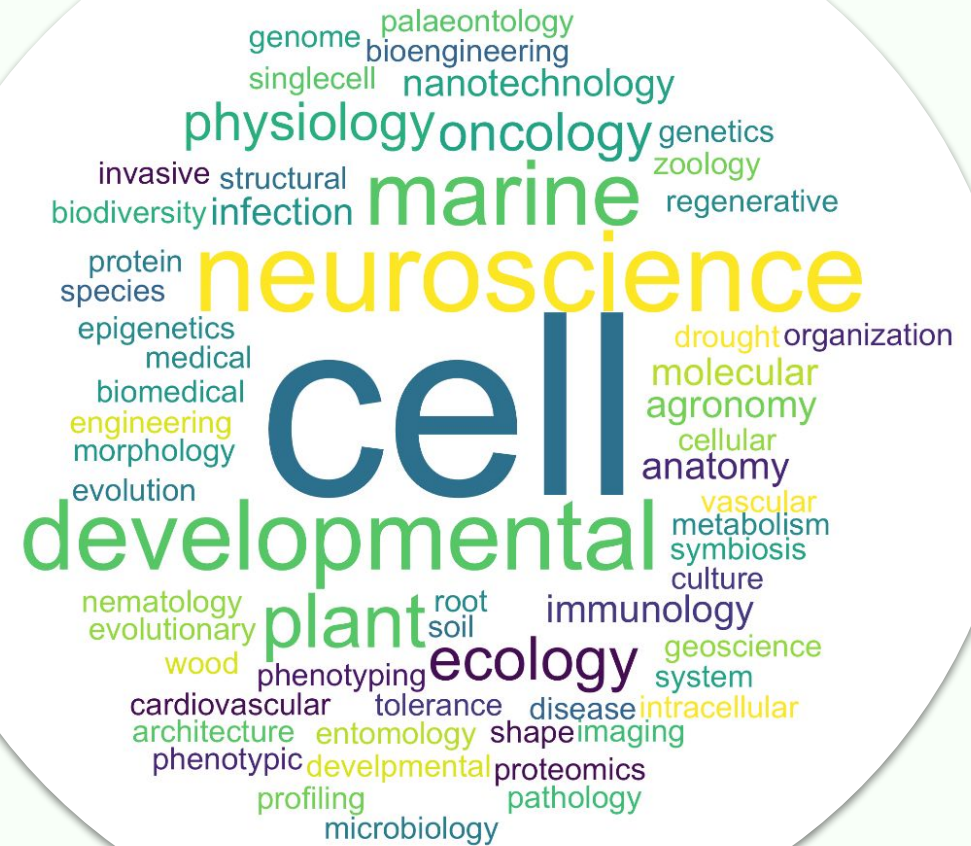
ai4life.eurobioimaging.eu

Evaluation & Timeline



Outcomes of the First Open Call

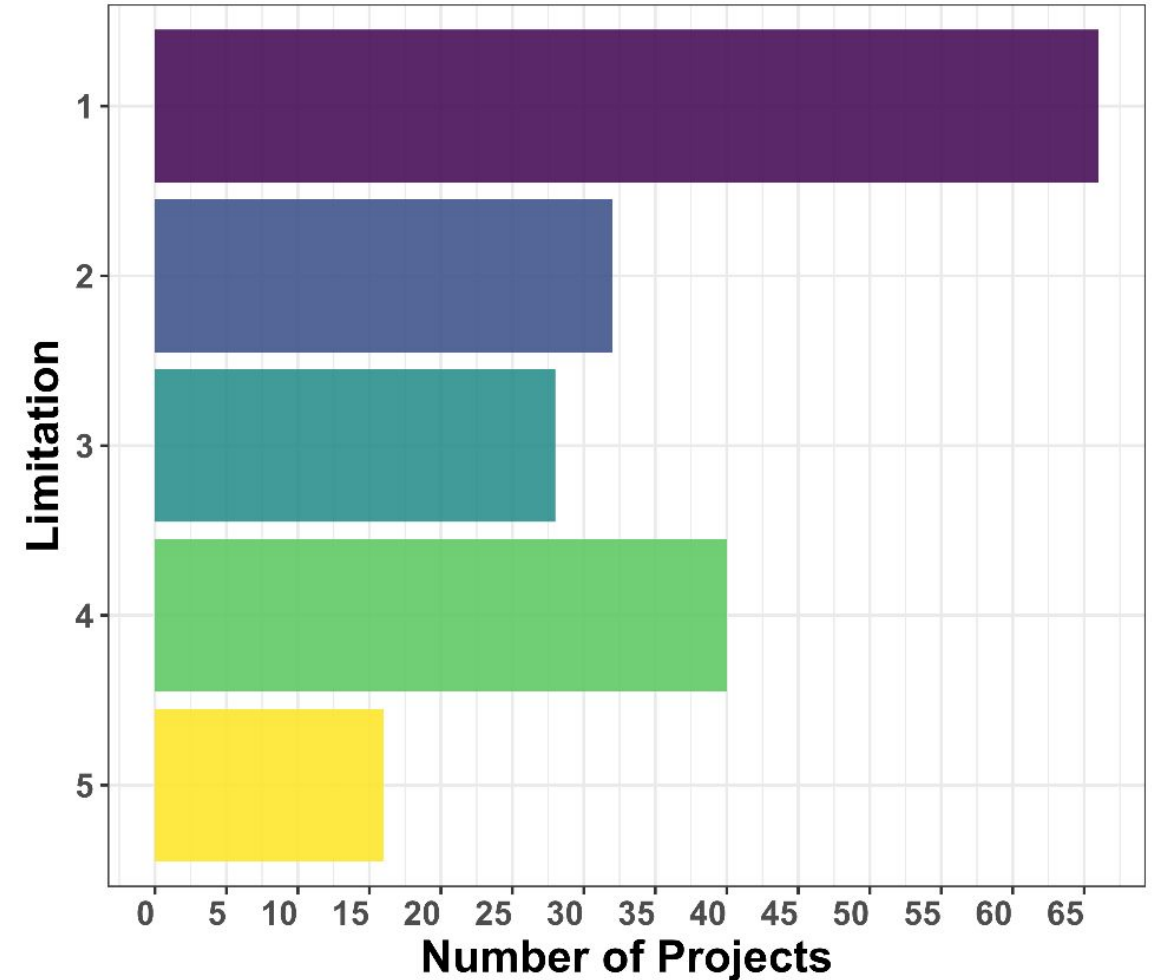
72 applications



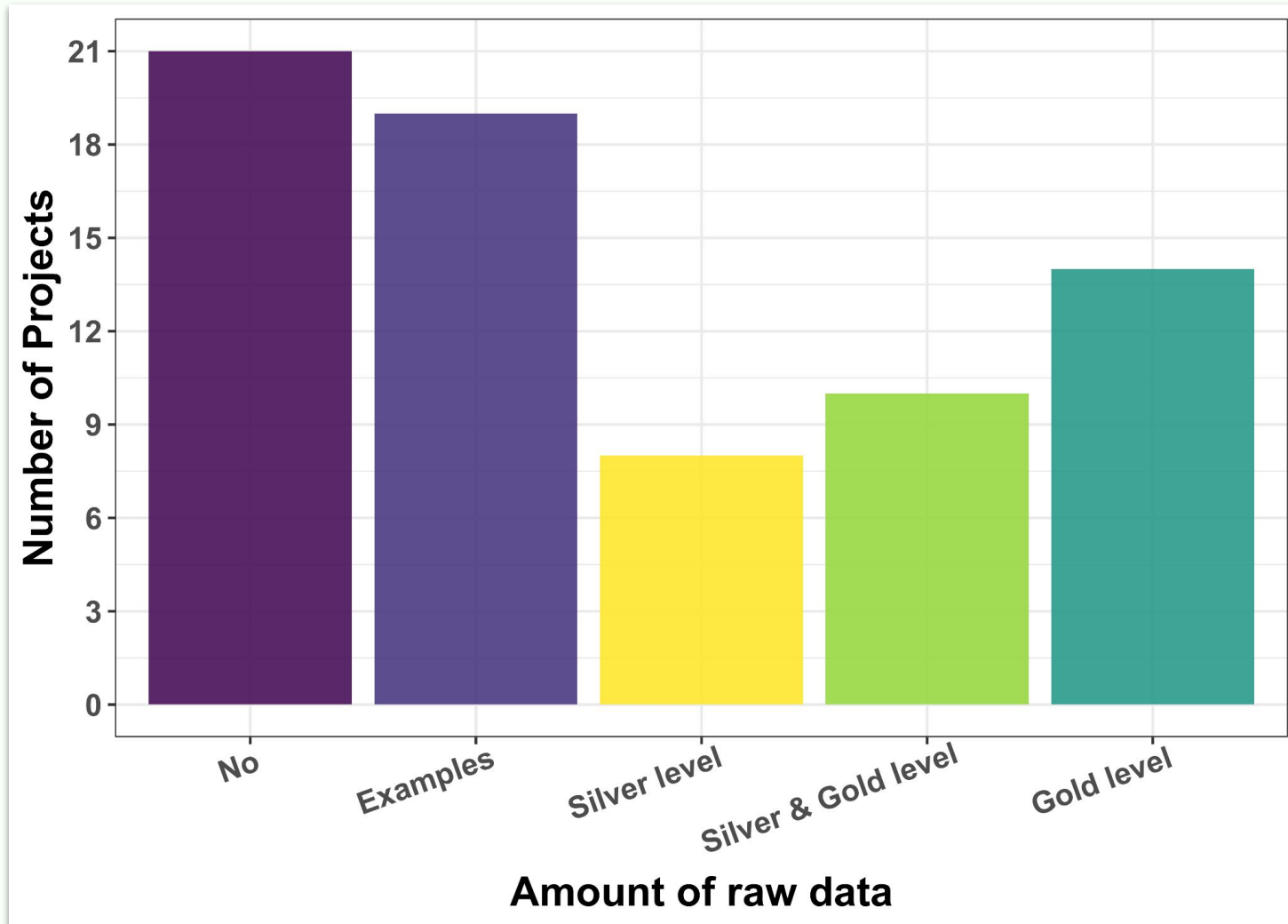
Why did they apply? What challenge are they facing?

Limitation in analyzing data

- Example 1: Improving your image analysis workflow
- Example 2: Improving your image analysis data and/or data storage strategy
- Example 3: Training data creation
- Example 4: Consultancy on available tools and solutions
- Example 5: Catering your problem to other experts



Labeled data available



Silver ground truth

Results/labels good enough to be used for publication (but maybe fully or partly machine generated).

Gold ground truth

Human curated to further reduce remaining mistakes.

OPEN CALL

for Bioimage
Analysis Support

8
projects
awarded

- Analysis of the fiber profile of skeletal muscle.
- Atlas of Symbiotic partnerships in plankton revealed by 3D electron microscopy.
- Automated and integrated cilia profiling.
- Identifying senescent cells through fluorescent microscopy.
- Image-guided gating strategy for image-enabled cell sorting of phytoplankton.
- Leaf tracker plant species poof.
- SGEF, a RhoG-specific GEF, regulates lumen formation and collective cell migration in 3D epithelial cysts.
- Treat CKD.



AI4Life



ai4life@eurobioimaging.eu

ai4life.eurobioimaging.eu

<https://ai4life.eurobioimaging.eu/first-ai4life-open-call-announcement-of-selected-projects/>

Thank you!

- Community meetings every Wednesday at 4pm CE(S)T
- Events:
<https://ai4life.eurobioimaging.eu/events/>
- Subscribe to the AI4Life newsletter:
<https://bit.ly/ai4life-newsletter>
- Reach out:
beatriz.serrano.solano@eurobioimaging.eu