

NL-Bioimaging: A Distributed FAIR Infrastructure for BioImaging in the Netherlands

Dr Katy Wolstencroft, Leiden Institute of Advanced Computer Science



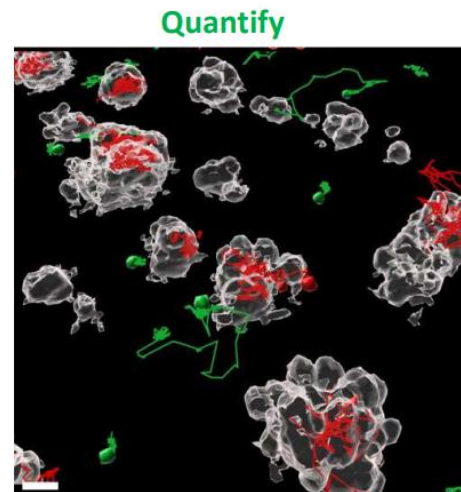
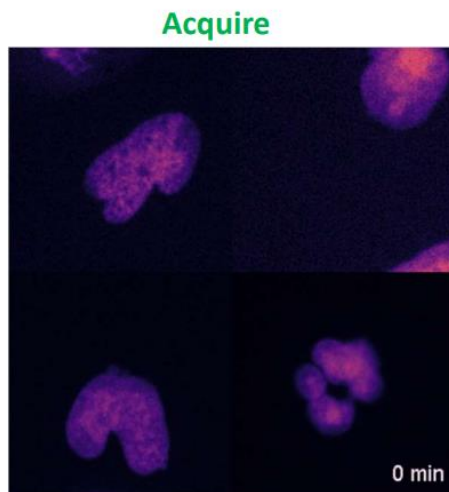
Universiteit
Leiden
The Netherlands



NL-BIOIMAGING AM

Aims and Challenges

- Open access to microscopy infrastructure from **18 institutes**
 - **Elevate bioimaging across the Netherlands**
- Bridging biological and technological scales
- Innovative microscopy analyses, self-steering, AI driven
- **National infrastructure for sharing FAIR data and image analyses**

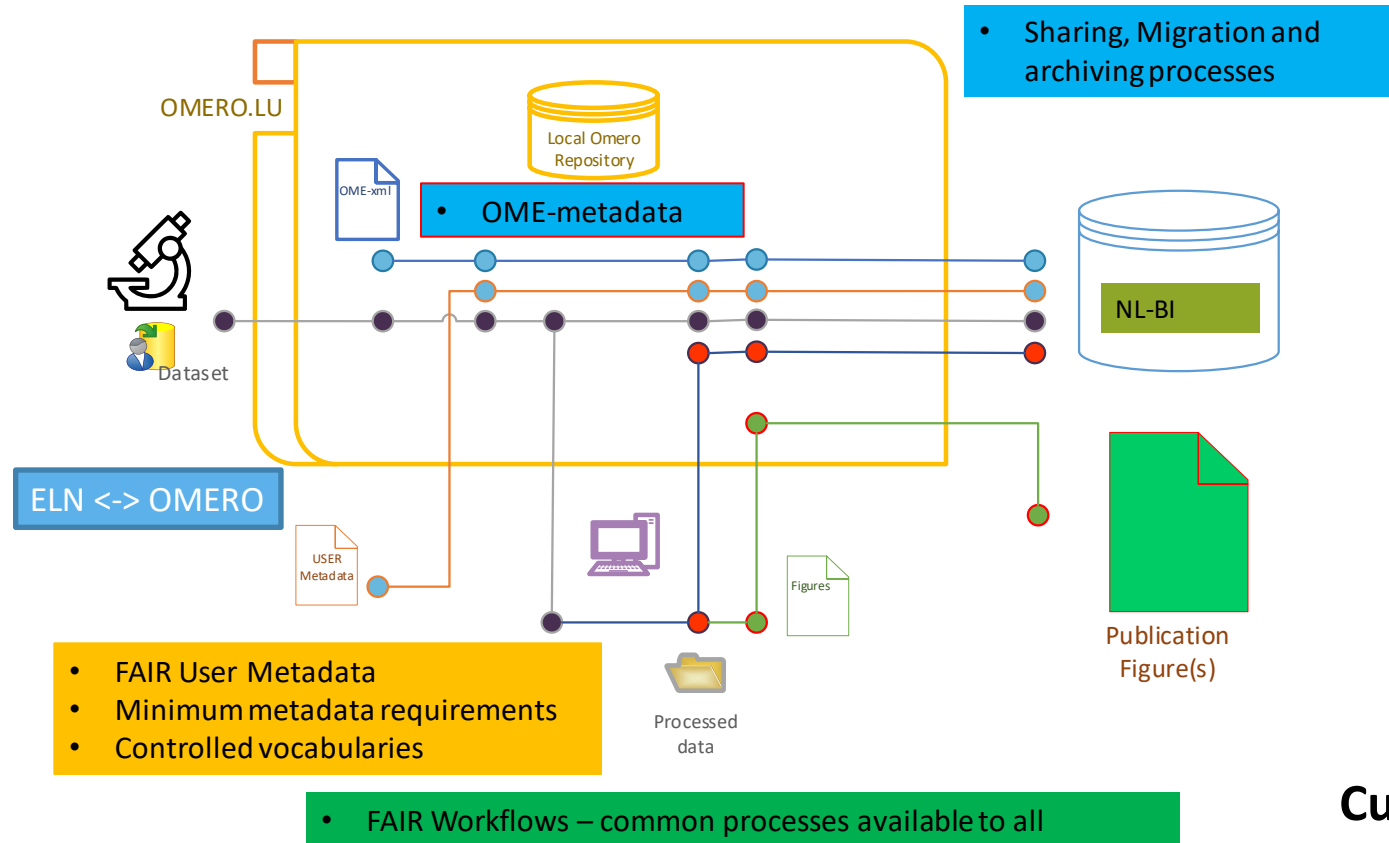


First Major Job: Standardise Local RDM

- Data generated and stored at each node
 - Distributed, local data stores
- Different resources for RDM
- Work with local infrastructure
 - iRODS/OMERO/both/**nothing**
- **Propose minimum NL-BI standard**
 - **OMERO**
- **Improve the FAIRness of data lifecycle**
- **Enable collaboration across nodes**



Standardising Local RDM – Leiden FAIR Cell Observatory



Pilot Instance Sept 2020

- 4 microscopes
- 60 active users

Full Roll-out Feb 2021

- 50+ microscopes
- 104 active users

Current usage

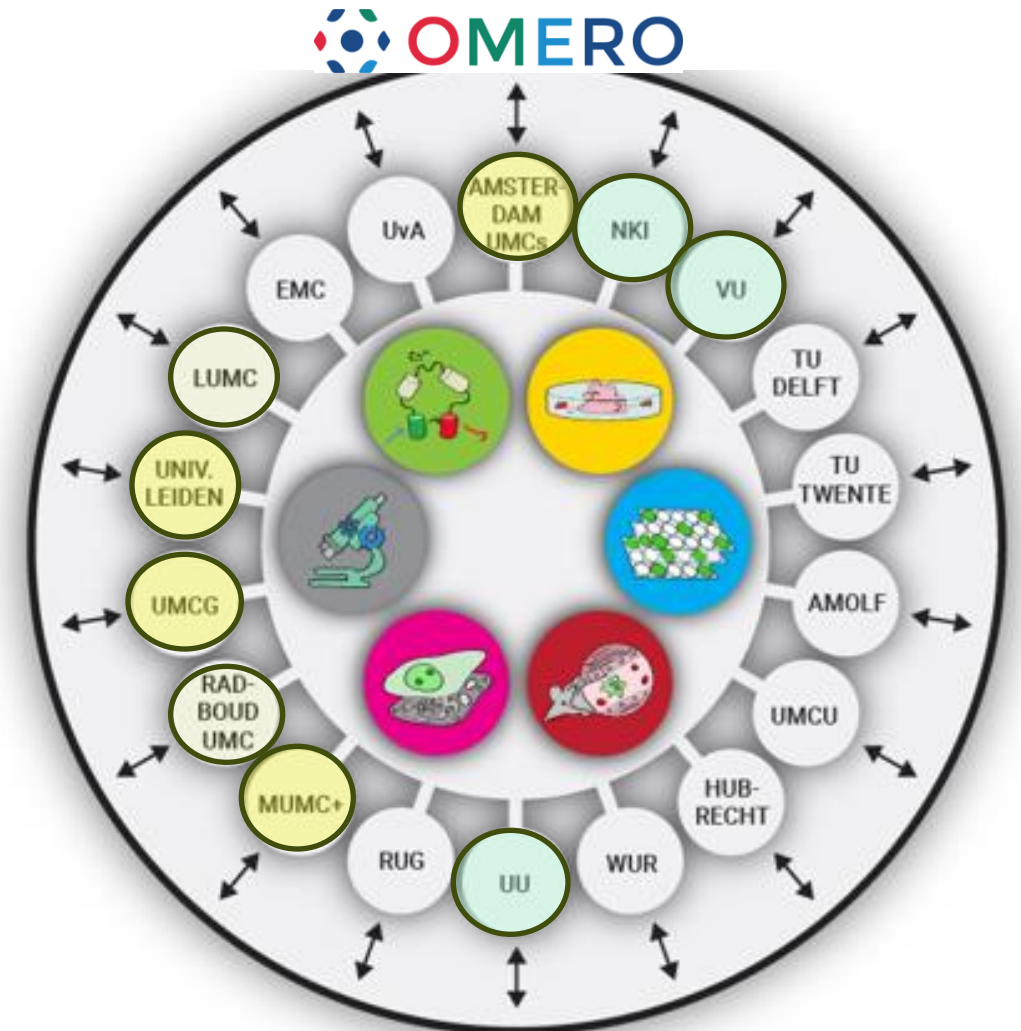
- 60+ microscopes
- 400+ active users

Training, documentation and support
.....still discussing cost model and long-term support

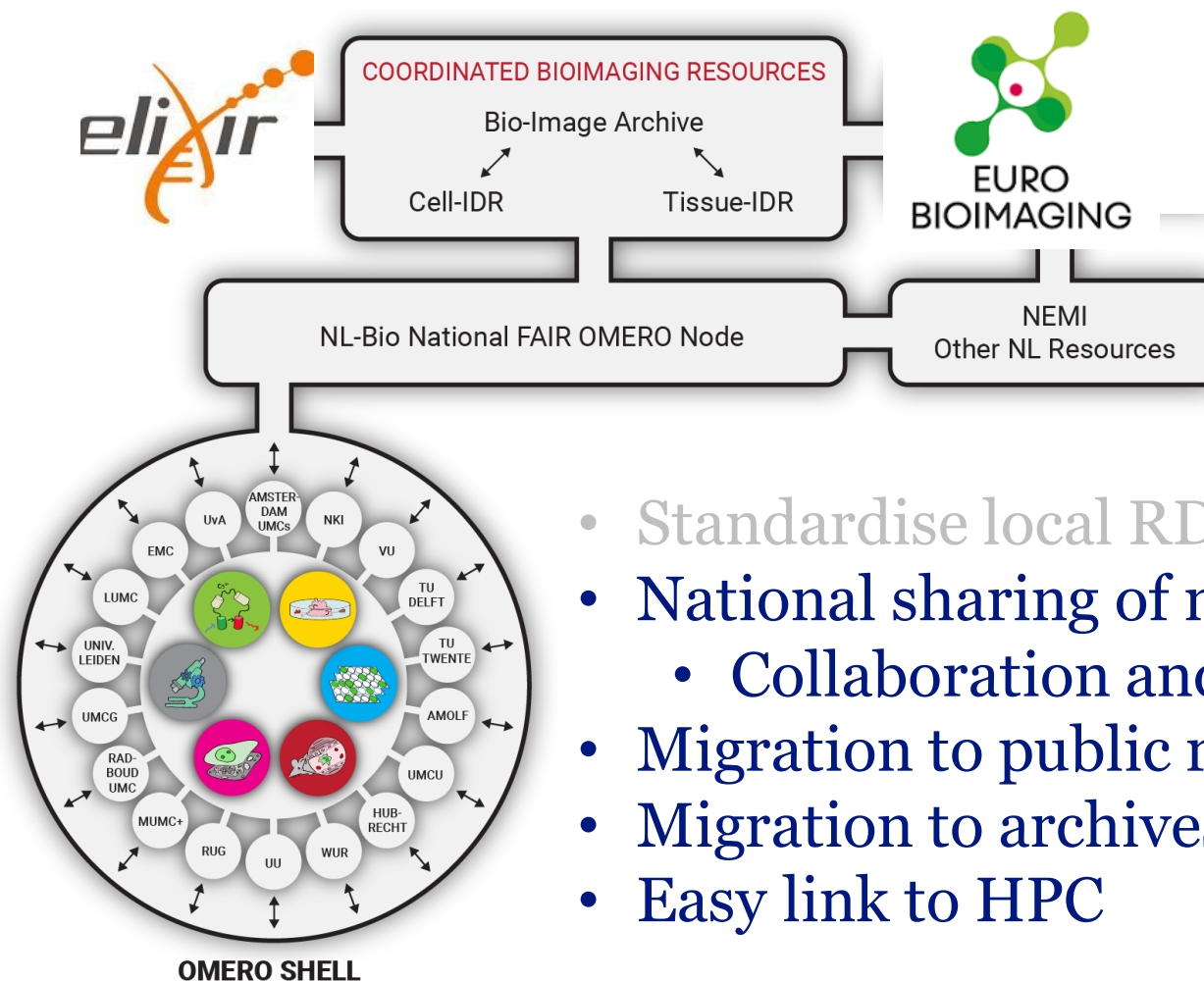
Rolling Out Nationally

- National network for DM expertise
- Sharing training/documentation/time
- OMERO established in multiple places
- iRODS in other places
- **SURF SRAM for access across instances**

- **NL-Bioimaging OMERO docker image**
 - Sync new versions
 - Roll out new features
 - Shared knowledge experience



The National Infrastructure and Beyond

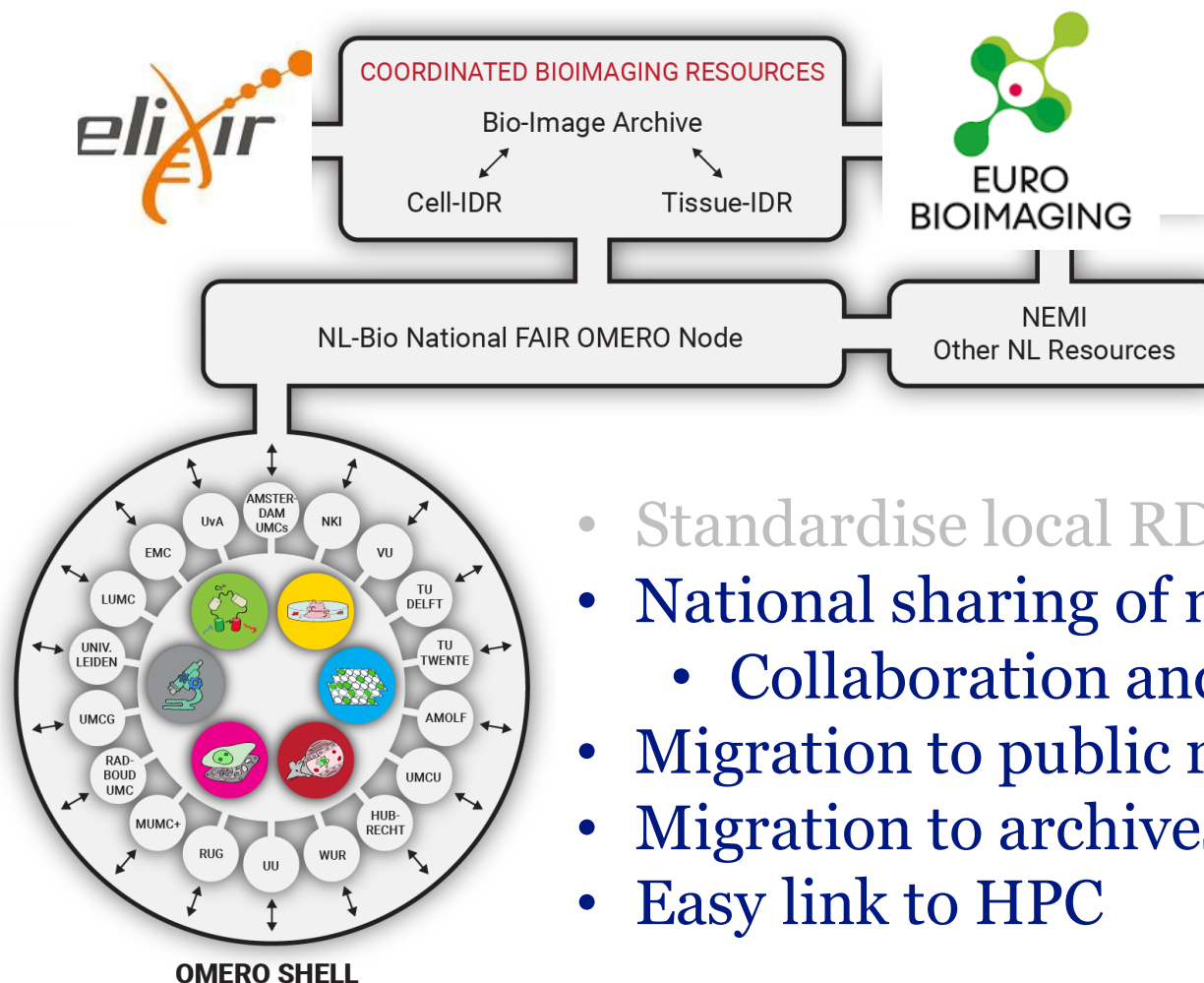


- Standardise local RDM
- National sharing of metadata
 - Collaboration and dissemination
- Migration to public repositories
- Migration to archives – cost reduction
- Easy link to HPC

The National Infrastructure and Beyond

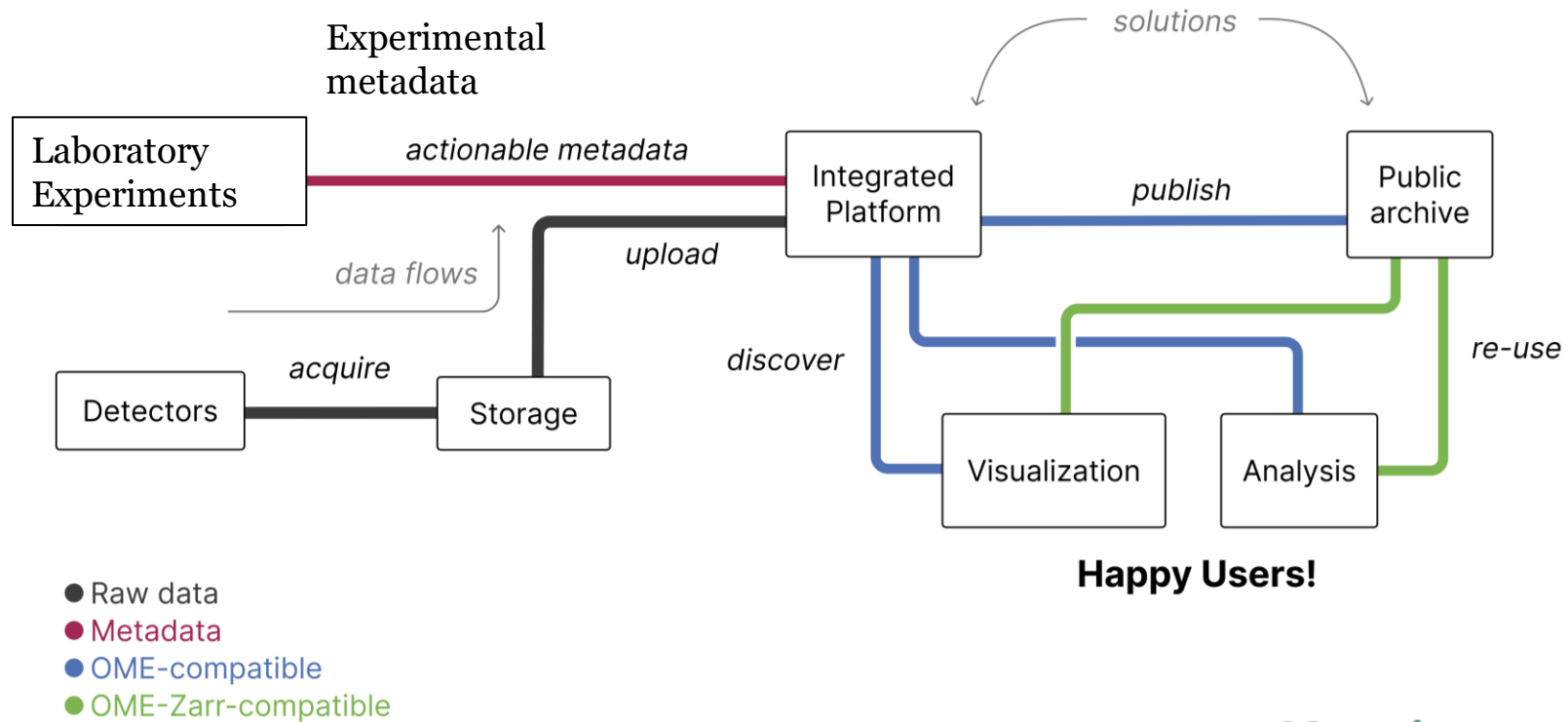
Necessary Components

- Standards – metadata, data and analyses
- Packaging – FAIR Digital Objects
- **Common Data Sharing Policy**



- Standardise local RDM
- National sharing of metadata
 - Collaboration and dissemination
- Migration to public repositories
- Migration to archives – cost reduction
- Easy link to HPC

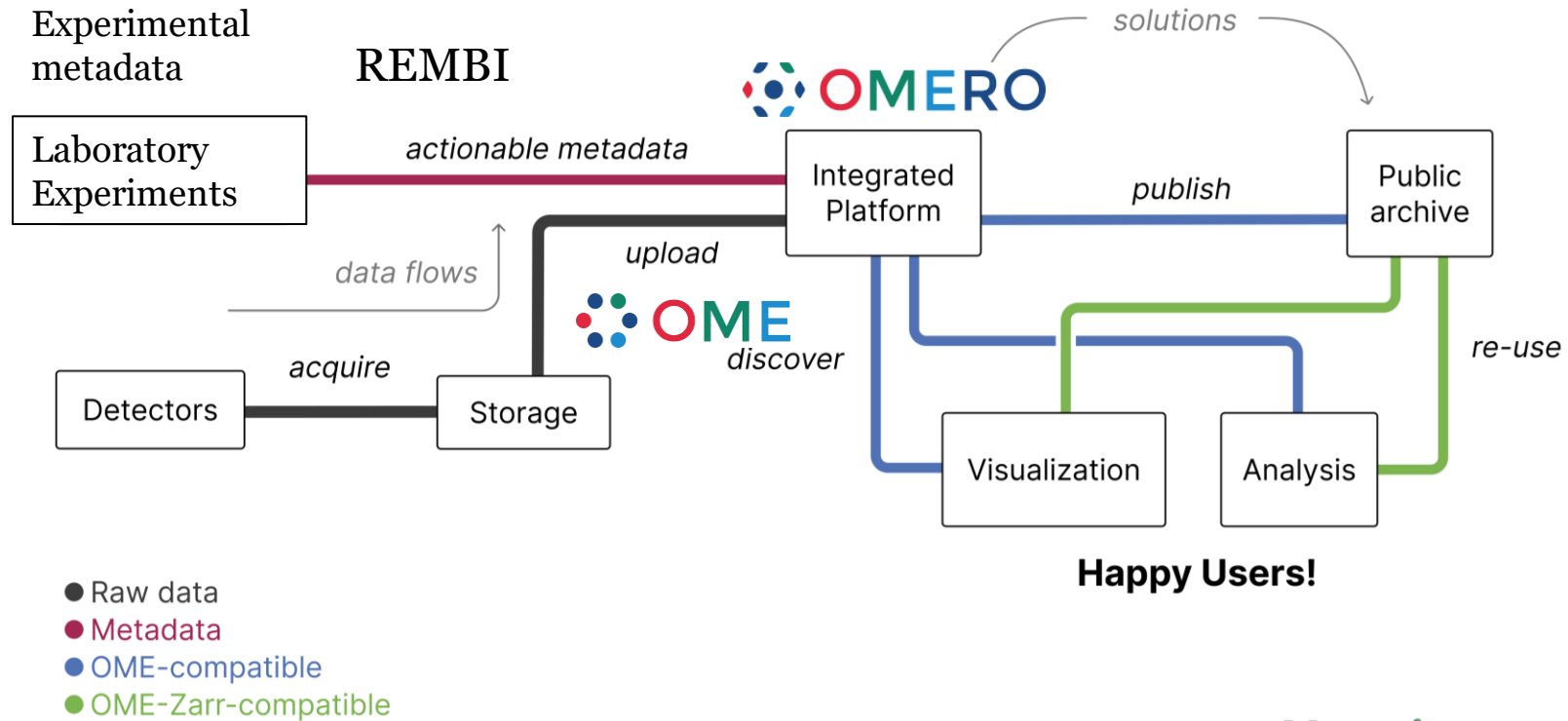
What we are aiming for



Happy Users!

Utopian
Biolmage Town (BIT)
FAIR Data Metro Map

What we are aiming for

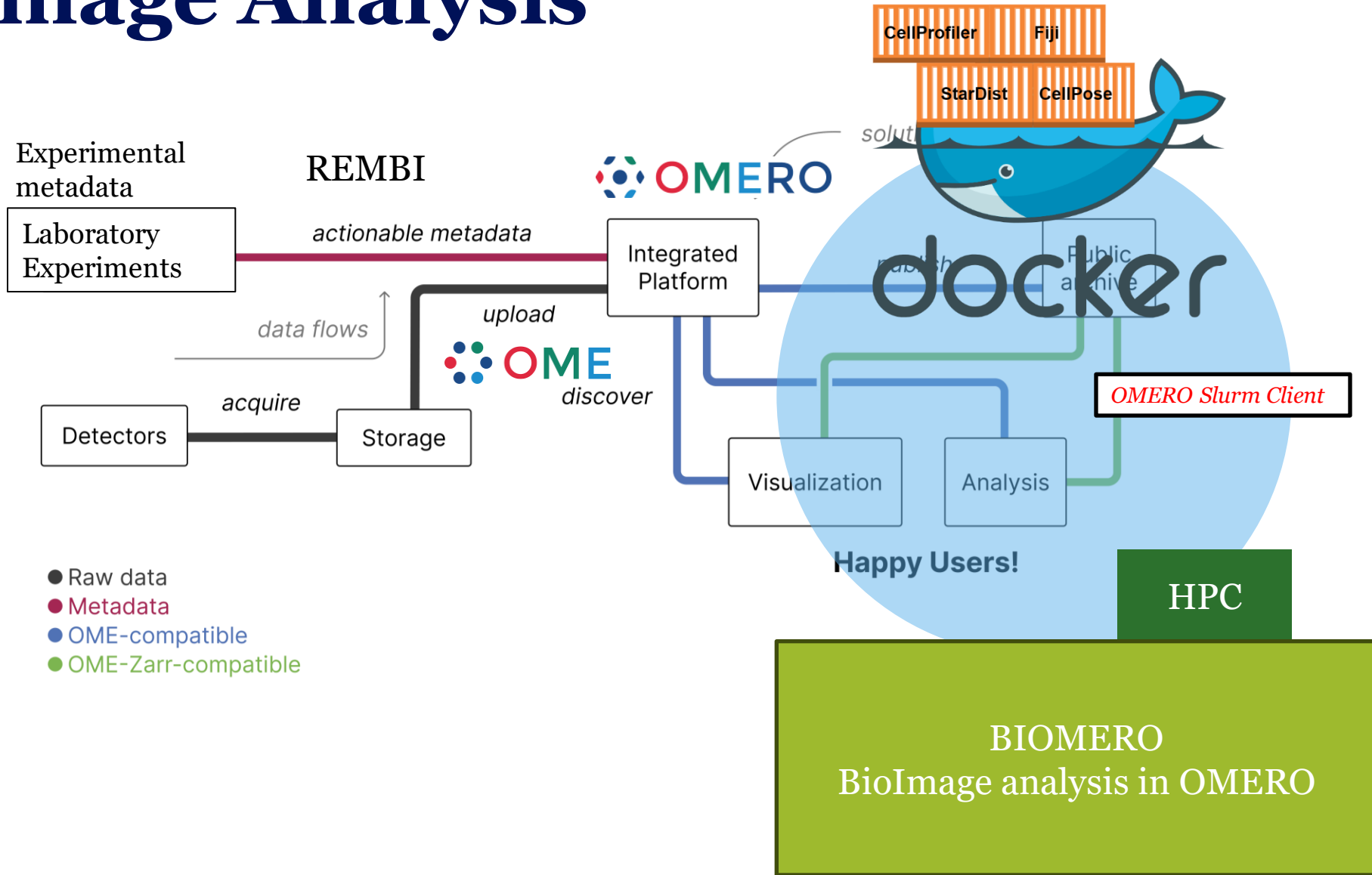


Happy Users!

Utopian
Biolmage Town (BIT)
FAIR Data Metro Map

FAIR Image Analysis

F A I R
 Findable Accessible Interoperable Reusable



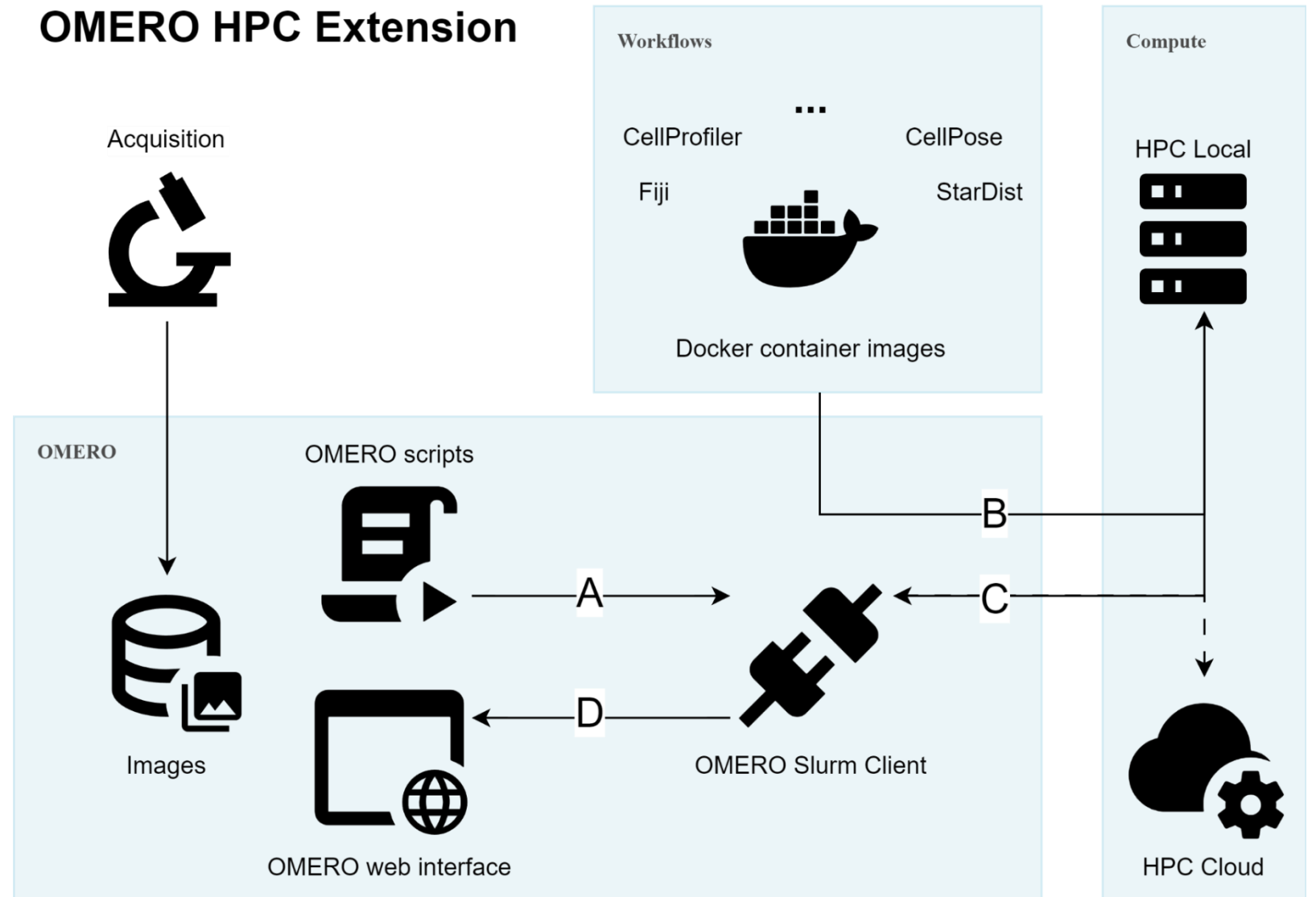
OMERO + SLURM HPC

Easier computational workflows, using OMERO data on HPC

Based on Bioflows/cytomine

- Containerised (environment)
- Metadata description

OMERO HPC Extension



<https://arxiv.org/abs/2402.00734> Torek Luik et al Amsterdam UMC



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



Rich metadata and data findable by both humans and computers

Formal language for knowledge representation

Protocols and policies for data (and metadata) access and use

Plurality of accurate and relevant attributes
Detailed provenance
License

Data
Models
Methods/protocols
Software
Research Assets

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



Metadata in community standard formats with agreed vocabularies

Who can see my data and in what circumstances?

Links between data
Which standards
License

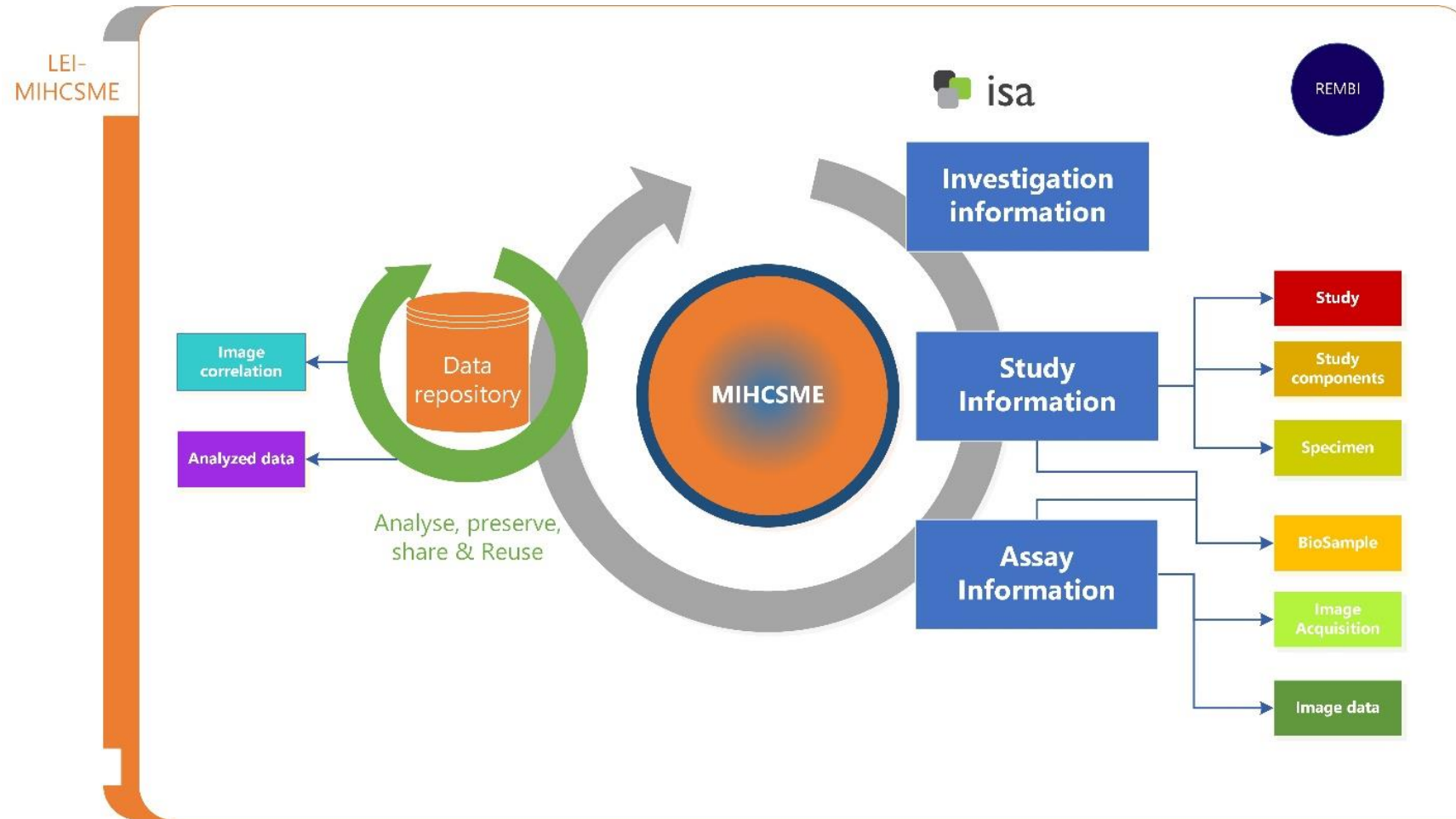
Data
Models
Methods/protocols
Software
Research Assets

Standards for Supporting the Data Life Cycle



- DM is not just for publication
- Consider data from experimental design onwards
- **General guidelines for all microscopy**
 - Specification for sub-domains?
- **What if it's not only microscopy data?**
 - How do we interlink with other data?

Linking REMBI and ISA for Sub-Domains



- Easier integration with other data sets (e.g. Omics)
- Visual omics
- Different templates for different imaging modalities
- MIZME
- + more

MIHCSME – High Content Screening

Selected: Minimal Information for High Content Screening in Microscopy Experiments (MIHCSME) (Investigation)

Tree Split Graph

Description: User metadata is an essential part of experimental data. Scientists need to understand underlying conditions and...

Fullscreen

SEEK ID: <https://fairdomhub.org/investigations/575>

- Minimal Information for High Content Screening in Microscopy Experiments (MIHCSME)
 - MIHCSME templates
 - General MIHCSME template
 - LEI-MIHCSME - empty template
 - LEI-MIHCSME Template guidelines
 - MIHCSME template example for "Temporal single cell cellular stress response activity toward toxicants inducing adaptive stress resp

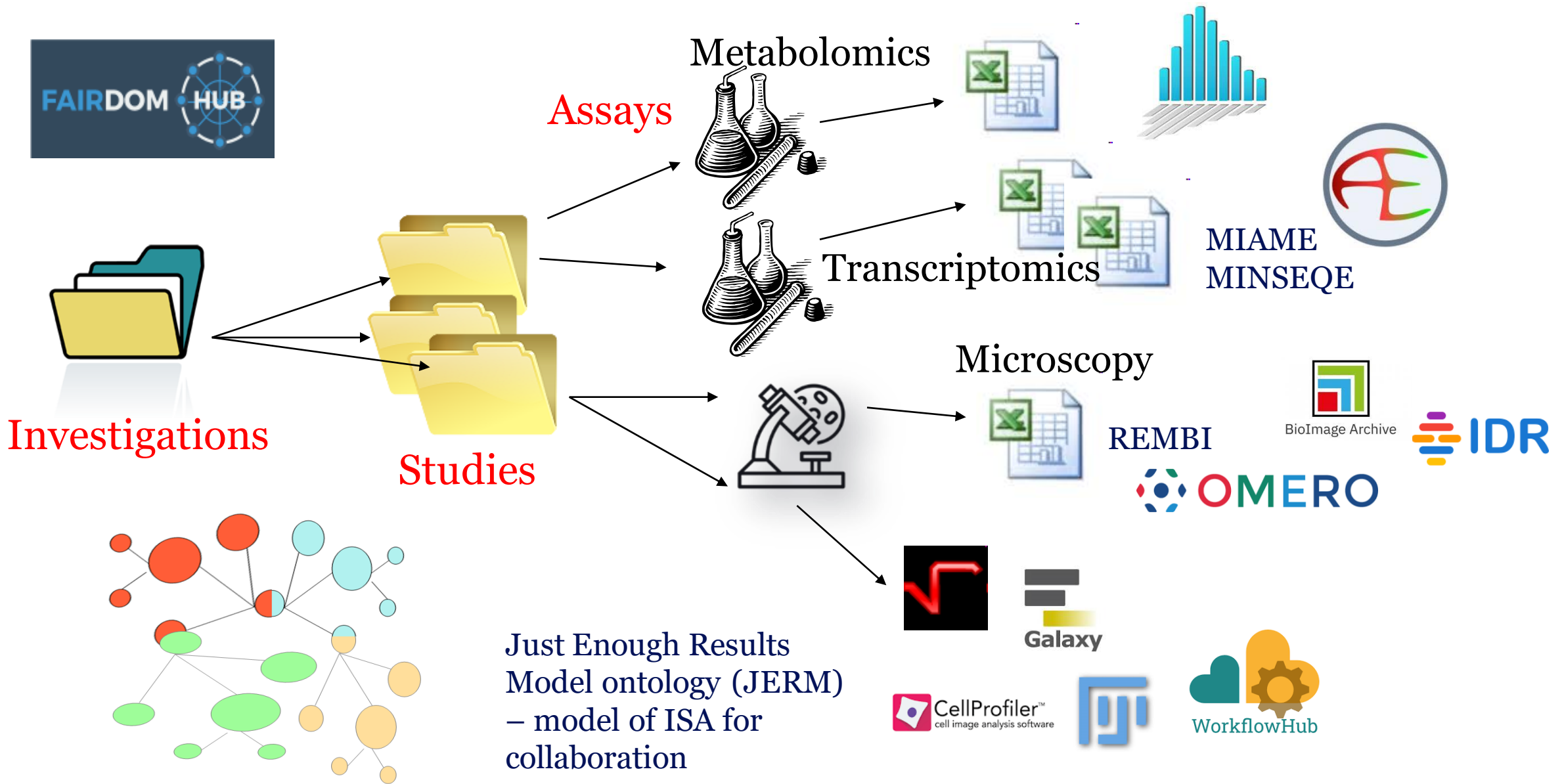
Hosseini, R., Vlasveld, M., Willemsse, J. *et al.* FAIR High Content Screening in Bioimaging. *Sci Data* **10**, 462 (2023)

Key	Value	Comments_User
Assay Title	HepG2 A20 GFP	
Assay internal ID	aLW113	
Assay Description	The screened was done in 384 well plate without randomization. Each plate contained only	
Assay number of biological replicates		3
Number of plates		3
Assay Technology Type	confocal microscopy	
Assay Type	high content screen of cells in treated with a compound library	
Assay External URL	https://www.ebi.ac.uk/biostudies/eu-toxrisk/studies/S-TOXR1741	
Assay data URL	T:\CellObservatory\LACDR\DDS\DDS_2\Bob van de Water_SysBioTop_TQ	
Imaging protocol	The plates were imaged under the confocal microscope. The plates were imaged once every	
Sample preparation protocol	Live cells	
Cell lines storage location	Freezer -80 - tower D4	mother stock
Cell lines clone number	clone 21	Different clone numbers used in this study. For
Cell lines Passage number	15-25	
Image number of pixelsX		512

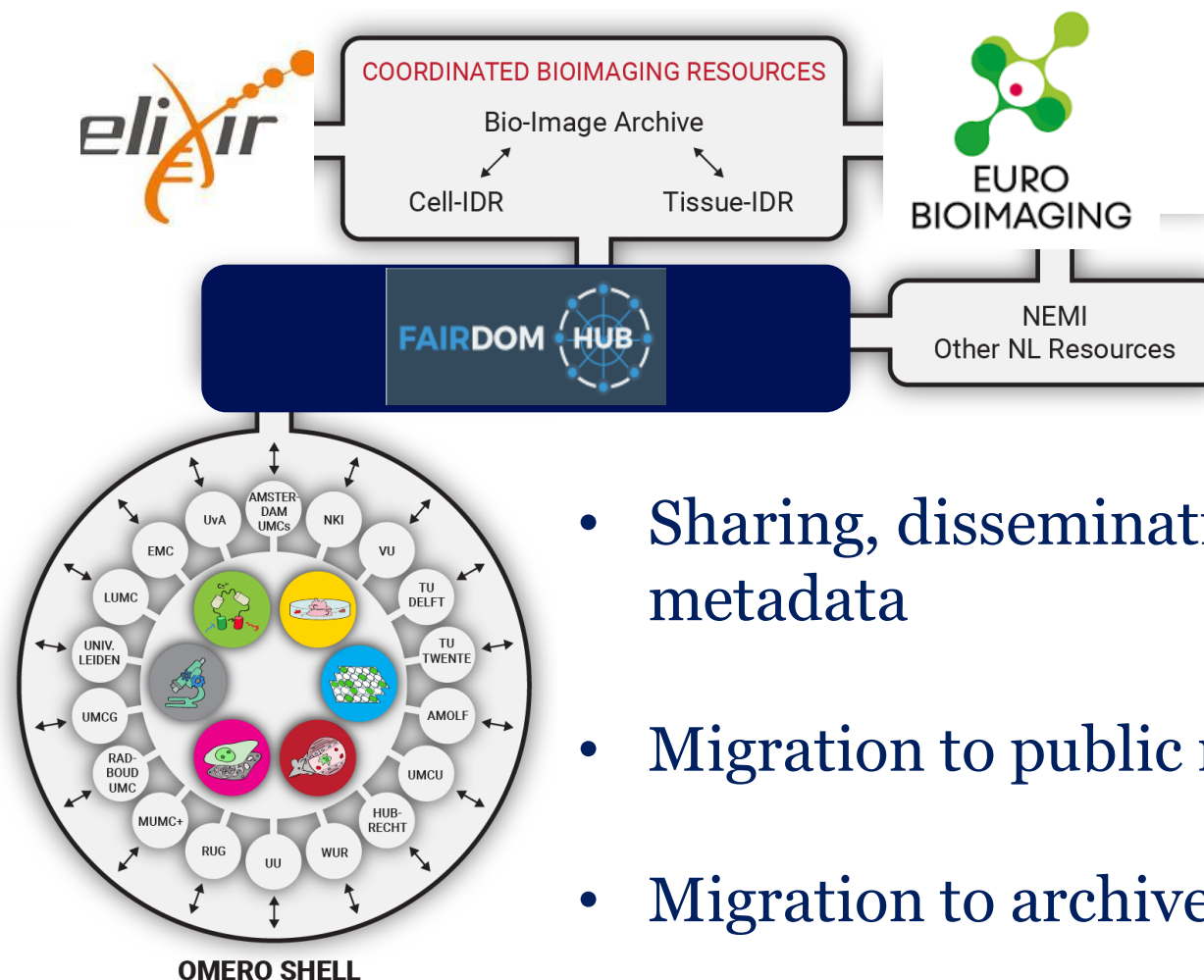
StudyInformation AssayInformation AssayConditions

SOP Library file

Spreadsheets to Semantic Models: JERM

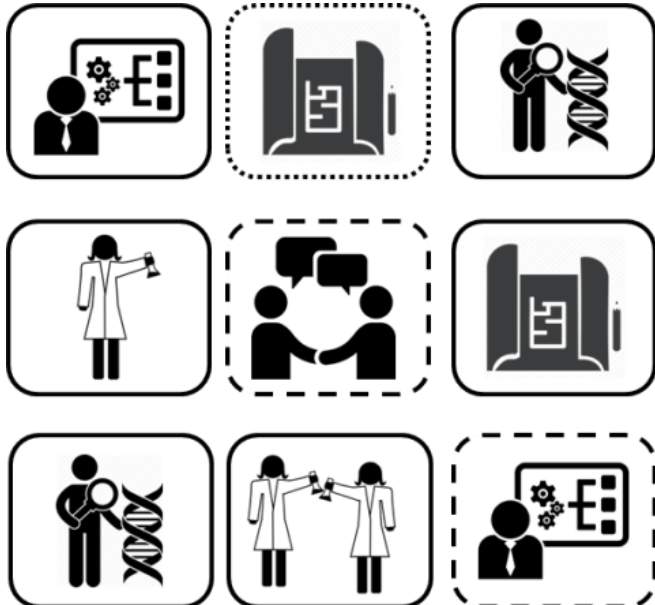
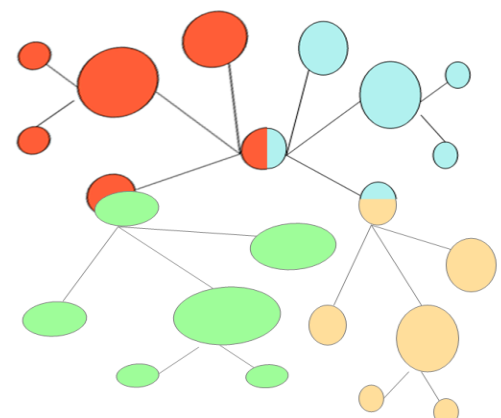


Requirements: NL-Bioimaging National Node



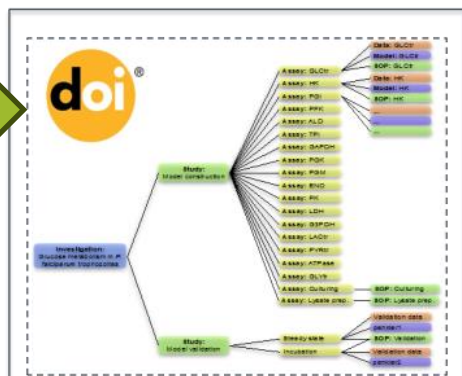
- Sharing, dissemination and linking of metadata
- Migration to public repositories
- Migration to archives

Packaging and Linking with FAIRDOM



- Asset
- Equipment
- Instrument
- Software
- Experiment_description
- Assay
- Investigation
- Study
- Experimental_factors
- Assay_type
- Bio_material
- Sample
- Specimen
- Characteristic_for_specimen_or_sample
- Factors_studied
- Measured_item
- Biological_entity
- Microarray_factors
- Proteomics_factors
- Ionization_source
- Normalization_method
- Post_source_component
- Quantitation
- Spectrum_generation

Just
Enough
Results
Model



My Investigation
Sept 2023

A project commons
Self-managed spaces
Yellow pages

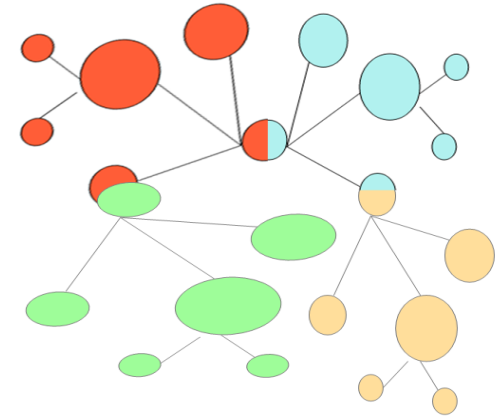
**RDF as a Bridge to Domain Platforms,
like OMERO, Or there and back again
SWAT4HCLS 2024, Moore et al**

Moving Forward

- Practical, pragmatic progress – demonstrated benefits of OMERO and FAIR approach

Moving Forward

- OMERO and iRODS integration
- FAIRDOM - IRODS
- OMERO and OME-NGFF
- REMBI compliant metadata – automation for submission to public archives
- Domain ontologies – EDAM-Bioimaging, FBBI, OME etc



Acknowledgements



- NL-Bioimaging Data Management and Analysis National team
- Rohola Hosseini (OMERO admin and Rolling-out national OMERO for NL-BI)
- Joost Willemse (OMERO admin IBL)
- Sylvia Le Dévédec (OMERO admin and HCS lead)
- Robert Schultz (OMERO server admin)
- Lennard Voortman (DM Analysis Coordinator)
- Pascal de Boer (OMERO Groningen and iRODS)
- Rodrigo Rosas (Containerised OMERO)
- Torec Luik (BIOMERO)
- Przemek Krawczyk (BIOMERO)
- Ron Hoeb (BIOMERO)
- Ben Giepmans – NEMI and NL-BI co-lead
- Eric Reits – NL-Bioimaging co-Ordinator

