



Multi-Modal Data Integration and Machine Readability between OMERO and ARC

Mohamed M. Abdrabbou¹  <https://orcid.org/0000-0003-2778-7483>, Mehrnaz Babaki^{2,3} 
<https://orcid.org/0000-0003-3099-618X>, Tom Boissonnet⁴  <https://orcid.org/0000-0002-3328-9467>, Michele Bortolomeazzi⁵
 <https://orcid.org/0000-0001-5805-5774>, Jean-Marie Burel⁶  <https://orcid.org/0000-0002-1789-1861>, Eik Dahms⁷ 
<https://orcid.org/0000-0002-7452-1146>, Vanessa Fuchs⁴  <https://orcid.org/0000-0002-4101-6987>, Niraj Kandpal² 
<https://orcid.org/0009-0007-5101-4786>, Christoph Möhl²  <https://orcid.org/0000-0002-0829-5101>, Josh Moore⁸ 
<https://orcid.org/0000-0003-4028-811X>, Timo Mülhhaus⁹  <https://orcid.org/0000-0003-3925-6778>, Astrid Schauss^{2,3} 
<https://orcid.org/0000-0002-6658-2192>, Kevin Schneider⁹  <https://orcid.org/0000-0002-2198-5262>, Andrea Schrader¹⁰ 
<https://orcid.org/0000-0002-3879-7057>, Torsten Stöter¹¹  <https://orcid.org/0009-0003-6932-023X>, Monica
Valencia-Schneider⁷  <https://orcid.org/0000-0003-3430-2683>, Stefanie Weidtkamp-Peters⁴ 
<https://orcid.org/0000-0001-7734-3771>, Heinrich Lukas Weil⁹  <https://orcid.org/0000-0003-1945-6342>, Peter Zentis² 
<https://orcid.org/0000-0002-6999-132X>

¹ Department of Biology/Chemistry and Center for Cellular Nanoanalytics, University Osnabrück, Germany

² CECAD Imaging Facility, University of Cologne, Cologne, Germany

³ Center for Molecular Medicine Cologne (CMMC), Cologne, German

⁴ Center for Advanced Imaging, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

⁵ Single-cell Open Lab, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁶ Divisions of Molecular Cell and Developmental Biology, and Computational Biology, University of Dundee, Dundee, Scotland, UK

⁷ IT Center University of Cologne (ITCC), Cologne, Germany

⁸ German BioImaging—Society for Microscopy and Image Analysis e.V., Constance, Germany

⁹ Computational Systems Biology, University of Kaiserslautern-Landau, Kaiserslautern, Germany

¹⁰ CEPLAS Data Science and Data Management, University of Cologne, Germany

¹¹ Leibniz Institute for Neurobiology (LIN), Magdeburg, Germany

Correspondence: Niraj Kandpal, nkandpa2@uni-koeln.de

Abstract

Bioimaging datasets are acquired in a variety of file formats that often demand specialized software to be managed according to RDM (Research Data Management) standards. OMERO [1] (Open Microscopy Environment Remote Objects), a (meta)data management platform, specialized in microscopic imaging, is designed for this purpose and is a widely utilized tool in microscopic imaging [2]. More generic approaches like ARC [3] (Annotated Research Context), based on FDO [4] (FAIR Digital Object) principles, provide data annotation for heterogeneous data types respecting their individual research context. By design, an ARC can handle diverse data by combining a hierarchical directory structure with

metadata annotation defined by the ISA (Investigation, Study, Assay) abstract model [5] but does not yet provide capabilities for working with complex microscopic images. Here, our work presents various approaches to integrate these technologies to make the bioimaging data management interoperable.

To make use of the respective strengths of OMERO and ARC, raw data, related metadata and analysis results are converted to be consumable by each. However, both OMERO and ARC have different folder structures and therefore, the OMERO-ARC interoperability requires mapping of respective folder structures. To this end, we developed a recommended mapping schema, workflows and a software plugin to transfer (meta)data from OMERO to ARC and ARC to OMERO. The omero-arc plugin [6] being developed is based on the community tools omero-cli-transfer [7] and ARC Commander [8] and the goal is to automate the workflow of (meta)data exchange between OMERO and ARC. This way, the multimodal (meta)data in an ARC can be transferred and visualized in the graphical representation in OMERO to which the bioimaging users are accustomed, while changes in OMERO will be archived and versioned in the PLANT dataHUB [9].

For smoother interoperability of metadata and machine readability, an additional layer of metadata exchange and interoperability between OMERO and ARC can be a valuable step towards the goal of Fair Digital Objects. To this end, we are exploring plugins like omero-rdf [10] that creates a stream of RDF (Resource Description Framework) triples for metadata in OMERO and outputs it in various formats including json-ld & RO-Crate [11]. On the other hand, metadata from ARC can already be exported as a json-ld file with semantically rich metadata using various tools it provides to export ARC as RO-Crate. An ongoing effort of standardizing the contextual information in the bioimaging domain including metadata from OMERO will make the metadata machine readable, actionable and interoperable.

Keywords: Bioimaging, OMERO, ARC, RDM, FDO, RDF, RO-Crate, PLANT dataHub

Resources

- OMERO-ARC plugin - <https://github.com/cmohl2013/omero-arc> "A software plugin that enables transfer of (meta)data from OMERO to ARC."
- Imaging ARC - <https://git.nfdi4plants.org/NKandpal/LarvaeDrosophila> "Representation of bioimaging data as an ARC."
- OMERO ARC Workflow - DOI [10.5281/zenodo.15225615](https://doi.org/10.5281/zenodo.15225615) "A tutorial with example workflow on (meta)data exchange between OMERO and ARC."

Author contributions

Conceptualization - Peter Zentis, Josh Moore, Timo Mühlhaus, Niraj Kandpal, Heinrich Lukas Weil, Michele Bortolomeazzi, Kevin Schneider, Christoph Möhl, Jean-Marie Burel, Torsten Stöter, Tom Boissonnet, Eik Dahms

Data curation - Andrea Schrader, Torsten Stöter, Vanessa Fuchs, Mehrnaz Babaki, Tom Boissonnet, Michele Bortolomeazzi, Niraj Kandpal, Peter Zentis

Funding acquisition - Stefanie Weidtkamp-Peters, Astrid Schauss, Peter Zentis, Monica Valencia-Schneider, Josh Moore

Software - Christoph Möhl, Niraj Kandpal, Michele Bortolomeazzi, Mohamed M. Abdrabbou, Tom Boissonnet, Peter Zentis, Josh Moore

Supervision - Peter Zentis, Astrid Schauss, Monica Valencia-Schneider, Torsten Stöter, Josh Moore, Stefanie Weidtkamp-Peters

Writing – original draft - Niraj Kandpal, Peter Zentis

Writing – review & editing - Josh Moore, Andrea Schrader, Eik Dahms, Michele Bortolomeazzi, Monica Valencia-Schneider

Competing interests

The authors declare that they have no competing interests.

Funding

This work is part of NFDI4BIOIMAGE funded by the German Research Foundation (DFG) within the framework of the NFDI – project number: 501864659 and supported by DataPLANT - project number : 442077441

Acknowledgement

We would like to acknowledge de.NBI for organizing and supporting our project at the 2nd biohackathon germany.

References

- [1] Chris Allan, Jean-Marie Burel, Josh Moore, Colin Blackburn, Melissa Linkert, Scott Loynton, Donald MacDonald, William J Moore, Carlos Neves, Andrew Patterson, Michael Porter, Aleksandra Tarkowska, Brian Loranger, Jerome Avondo, Ingvar Lagerstedt, Luca Lianas, Simone Leo, Katherine Hands, Ron T Hay, Ardan Patwardhan, Christoph Best, Gerard J Kleywegt, Gianluigi Zanetti & Jason R Swedlow (2012) OMERO: flexible, model-driven data management for experimental biology. *Nature Methods* 9, 245–253. Published: 28 February 2012, doi: [10.1038/nmeth.1896](https://doi.org/10.1038/nmeth.1896) PMID: [22373911](https://pubmed.ncbi.nlm.nih.gov/22373911/) PMCID: [PMC3437820](https://pubmed.ncbi.nlm.nih.gov/PMC3437820/)
- [2] Schmidt C, Hanne J, Moore J *et al.*, Research data management for bioimaging: the 2021 NFDI4BIOIMAGE community survey [version 2; peer review: 2 approved]. *F1000Research* 2022, **11**:638 doi: [10.12688/f1000research.121714.2](https://doi.org/10.12688/f1000research.121714.2)
- [3] DataPLANT community, ARC Specification 1.2 (2023). [Online]. Available: <https://doi.org/10.5281/zenodo.10091038>
- [4] Anders, I. et al., FDO Forum FDO Requirement Specifications (2023). [Online]. Available: <https://zenodo.org/records/7782262>
- [5] Sansone, S.-A. et al., ISA Model and Serialization Specifications 1.0 (2016). [Online]. Available: <https://doi.org/10.5281/zenodo.163640>
- [6] *OMERO-ARC Plugin*. (2023) C. Möhl [Online]. Available: <https://github.com/cmohl2013/omero-arc>
- [7] E. Martins Ratamero et al., ome/omero-cli-transfer: 0.8.0 - new `--simple` option in `pack` (0.8.0) (2023). [Online]. Available (all versions): DOI : <https://doi.org/10.5281/zenodo.7573591>
- [8] *ARC Commander*. (2022) Weil, H. L. et al. [Online]. Available: <https://github.com/nfdi4plants/ARCCommander>
- [9] Weil, H.L., Schneider, K., Tschöpe, M., Bauer, J., Maus, O., Frey, K., Brilhaus, D., Martins Rodrigues, C., Doniparthi, G., Wetzels, F., Lukasczyk, J., Kranz, A., Grüning, B., Zimmer, D., Deßloch, S., von Suchodoletz, D., Usadel, B., Garth, C. and Mühlhaus, T. (2023), PLANTdataHUB: a collaborative platform for continuous FAIR data sharing in plant research. *Plant J*, 116: 974-988. <https://doi.org/10.1111/tpj.16474>
- [10] Josh Moore, jean-marie burel, Torsten Stöter, & Stefan. (2024). German-BiolImaging/omero-rdf: v0.6.1 (v0.6.1). Zenodo. <https://doi.org/10.5281/zenodo.14415961>
- [11] Peroni S, Soiland-Reyes S, Sefton P, et al. Packaging research artefacts with RO-Crate. *Data Science*. 2022;5(2):97-138. doi:10.3233/DS-210053