

Author contributions

J.R.: conceptualization, software development, validation, data curation, writing review and editing. T.P.: conceptualization, software development, validation. A.R.: conceptualization, software development, validation. P.M.L.: conceptualization, validation, writing review and editing. M.S.I.: conceptualization, validation. L.A.C.:

conceptualization, validation, writing review and editing. G.M.: conceptualization, validation. C.S.-D.-C.: conceptualization, software, validation, supervision, funding acquisition. M.A.S.: conceptualization, writing review and editing. C.M.B.: conceptualization, validation, writing original draft, writing review and editing, supervision, project administration, funding acquisition.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41592-021-01290-5>.

Peer review information *Nature Methods* thanks Jon Mulholland, Brian Slaughter and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



MDEmic: a metadata annotation tool to facilitate management of FAIR image data in the bioimaging community

To the Editor — Although today the majority of scientific data, including microscopy and imaging data, are available in digital format, a real benefit from easy sharing and reuse of digital data according to the FAIR (findability, accessibility, interoperability and reusability) principles¹ exists only if data are understandable and

unambiguously interpretable. Collecting and maintaining the relevant metadata is key to ensuring that data are reliable and reusable and can be found and accessed by the scientific community. Imaging data are usually extremely rich data files, as they report on various parameters in a multidimensional space and are acquired

with complex microscopy instruments. The metadata or data models are very diverse due to the wide range of, for example, modalities, scales, experimental setups and file formats. Therefore, the appropriate use of suitable standardized metadata and data models is a challenge^{2,3}. Accordingly, flexible tools for capturing a complete

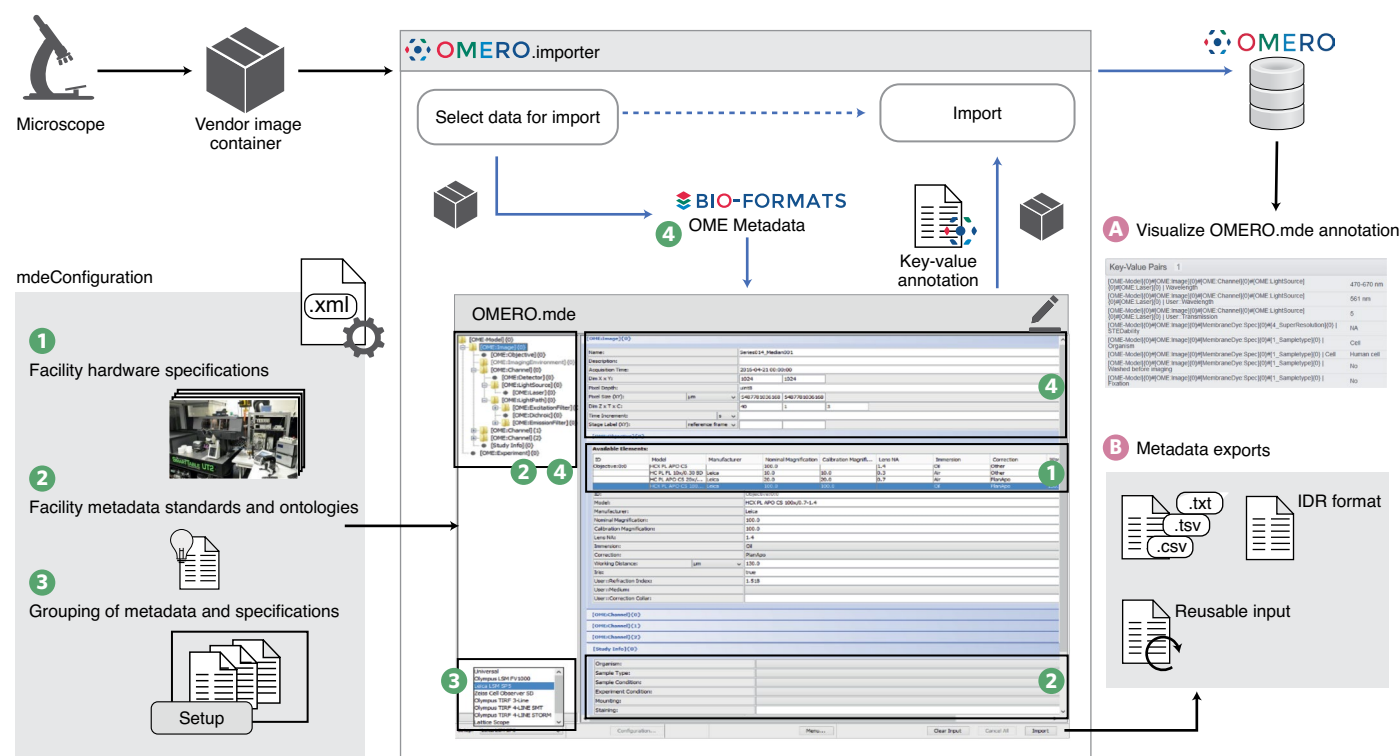


Fig. 1 | OMERO.importer with integrated MDEmic as OMERO.mde. In OMERO.importer, the MDEmic tool is integrated as an intermediate step for the selection of data for import and for the import itself. Metadata can then be added, which is transferred to the repository together with the image data (A); or the annotations can be exported in different formats in this step (B). MDEmic can be customized via a configuration file and loads the specifications from this file dynamically when the OMERO.importer is started (1, 2, 3). All technical metadata of the images marked in the previous step of data selection are read out by Bio-Format (4) and provided as values in MDEmic.

set of metadata are in great demand by researchers applying microscopy techniques. Moreover, it is important for imaging core facilities to be able to provide different standards with one tool and still be flexible enough to accommodate dynamic developments in methodology and analysis techniques. Many tools fail to strike this balance and therefore oscillate between using rigid data models and free text input without semantic context. Similarly, the integration and referencing of existing metadata is often lacking.

Our tool, MDEmic (MetaData Editor for microscopy; Supplementary Fig. 1), provides an easy and comfortable way to edit the metadata of microscopic imaging data and at the same time offers high flexibility in terms of the adjustment of metadata sets and their data models. As the standardization process in regard to the metadata of microscopic experiments is in full swing, MDEmic offers high flexibility for following this process. This is achieved through the dynamic configurability of both the queried and the integrated metadata and predefined values. The underlying data model can be extended sequentially by integrating these specified objects. The additional integration of ontology databases increases the interpretability of the data in general and ensures the discipline-specific integration of imaging data. MDEmic reads the technical metadata stored in the image file using Bio-Formats³, a software library for reading proprietary microscopy image (meta)data, and presents this metadata in the form of the OME (Open Microscopy Environment) Data Model^{4–6}. Visualizing this model as input forms allows the researcher to adjust or correct the technical metadata. In addition, based on the default in a configuration file associated with MDEmic, input forms for further metadata can be generated dynamically, integrating or extending the OME Data Model for technical metadata. The specification of metadata can include (i) type and category of metadata, (ii) fixed terms as selectable input values loading from subclasses by specification of ontology class identifiers and (iii) a defined structure for how to relate to other metadata categories. For all metadata, different sets of predefinitions can be integrated via the configuration file and selected by the user according to the respective scientific application or image technology. For example, fully described components of the existing microscope can be added or replaced by the user in the metadata collection. MDEmic is part of the standard installation package of the image database OMERO⁷ and is integrated in the OMERO.importer as OMERO.mde (Fig. 1). The OMERO.importer

can be used independent from a local OMERO instance by referring to a public OMERO instance such as the Image Data Repository, IDR^{8,9}. All metadata descriptions created in OMERO.mde can be saved and reloaded by the user for later reuse and adaptation or be exported to different textual formats. This functionality allows the output to be easily integrated with other needs. For example, this increases interoperability with other research data management tools to support integration with other data types or preparation of image data for publication or upload to public repositories¹⁰, such as the Image Data Resource (IDR; Supplementary Figs. 5 and 6).

In the following we describe a use case scenario that illustrates the direct benefits of MDEmic and OMERO.mde for researchers, wherein image data from samples treated with various membrane dyes are made available in a “Membrane Dye Database” hosted in the institutional OMERO instance and shared between the members of a collaborative research center. Here, we utilized the OMERO.mde extension for customized metadata annotation (Supplementary Figs. 2–4). For this purpose we have defined a metadata object called “Membrane Dye” in OMERO.mde. This object is available with different sub-objects describing the membrane dyes in more detail, such as “Effects On Sample” and “Internalization” of the dye. This object is provided together with the technical description and essential OME Data Model objects as input forms. All input forms are summarized in the setup “Membrane Dye Database.” This adjustment for the specific object “Membrane Dye” can be done by editing of the configuration file of OMERO.mde performed, for example, by a data steward of the CRC or a scientist of the imaging core facility.

Tools such as MDEmic, which significantly improve the interoperability of bioimaging data with other data types and within discipline-specific data management environments, are poised to have a high impact on the acceptance of FAIR research data management.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.






Data availability

All data are available under links listed in Supplementary Table 1.

Code availability

All code is accessible at Github (<https://github.com/ome/omero-insight>). The

fully open-source code link is listed in Supplementary Table 1. □

Susanne Kunis ^{1,6}, Sebastian Hänisch², Christian Schmidt ³, Frances Wong ⁴, Caterina Strambio-De-Castilla ⁵ and Stefanie Weidtkamp-Peters ^{2,6} ✉

¹Department of Biology/Chemistry, Centre for Cellular Nanoanalytics, University Osnabrueck, Osnabrueck, Germany. ²Centre for Advanced Imaging, University Duesseldorf, Duesseldorf, Germany. ³Bioimaging Centre, Department of Biology, University of Konstanz, Konstanz, Germany. ⁴Division of Computational Biology, Centre for Gene Regulation and Expression, University of Dundee, Dundee, UK. ⁵Program in Molecular Medicine, University of Massachusetts Chan Medical School, Worcester, MA, USA. ⁶These authors contributed equally: Susanne Kunis, Stefanie Weidtkamp-Peters. ✉e-mail: sukunis@uos.de; stefanie.weidtkamp-peters@hhu.de

Published online: 11 October 2021
<https://doi.org/10.1038/s41592-021-01288-z>

References

- Wilkinson, M. D. et al. *Sci. Data* **3**, 160018 (2016).
- Sarkans, U. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-021-01166-8> (2021).
- Linkert, M. et al. *J. Cell Biol.* **189**, 777–782 (2010).
- Goldberg, I. G. et al. *Genome Biol.* **6**, R47 (2005).
- Li, S. et al. *Methods* **96**, 27–32 (2016).
- Swedlow, J. R., Goldberg, I., Brauner, E. & Sorger, P. K. *Science* **300**, 100–102 (2003).
- Allan, C. et al. *Nat. Methods* **9**, 245–253 (2012).
- Ellenberg, J. et al. *Nat. Methods* **15**, 849–854 (2018).
- Williams, E. et al. *Nat. Methods* **14**, 775–781 (2017).
- Swedlow, J. R. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-021-01113-7> (2021).

Acknowledgements

We would like to thank the OME team for constant support and helpful discussion, in particular Josh Moore, Dominik Lindner and Jean-Marie Burel. We received funding from the Deutsche Forschungsgemeinschaft, grants SFB 1208 (ref: 267205415 INF/Z02, S.W.-P.) and SFB 944 (ref: 180879236 INF, S.K.); the Wellcome Trust (ref: 212962/Z/18/Z); the UK Biotechnology and Biological Sciences Research Council (ref: BB/R015384/1, F.W.); the US National Institutes of Health (ref: 2U01CA200059-06) and the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation, as part of their Imaging Scientist Program (ref: 2019-198155 (5022), C.S.D.C.).

Author contributions

S.K.: Software, writing—original draft, visualization; S.H.: visualization, investigation; C.S.: writing—review and editing; F.W.: writing—review and editing; C.S.D.C.: writing—review and editing; S.W.-P.: conceptualization, writing—original draft.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41592-021-01288-z>.