OME Files - An open source reference library for the OME-XML metadata model and the OME-TIFF file format

Roger Leigh[1], David Gault[1], Melissa Linkert[2], Jean-Marie Burel[1], Josh Moore[1], Sébastien Besson[1], Jason R. Swedlow[1,2*]

[1]Centre for Gene Regulation & Expression and Division of Computational Biology, University of Dundee, Dundee, Scotland, UK

[2]Glencoe Software, Inc. Seattle, WA, USA

*Correspondence:
email: jrswedlow@dundee.ac.uk
phone +44 1382 385819
fax +44 1382 388072
URL: www.openmicroscopy.org

Digital imaging is now used throughout biological and biomedical research to measure the architecture, composition and dynamics of cells, tissues and organisms. The many different imaging technologies create data in many different formats. This diversity of data formats arises because the creation of common, cross-domain data standards is difficult and the pace of innovation challenges the durability of any defined standard and the implementation of data stewardship principles such as FAIR (*1*). The Open Microscopy Environment (OME; http://openmicroscopy.org) has therefore developed the OME Data Model, an open, extensible specification for imaging metadata, that supports metadata related to an imaging experiment, data acquisition and any derived analytic results (*2-4*). The OME Data Model has been implemented in OME-XML so that metadata can be stored and accessed by any software. OME-TIFF embeds OME-XML into the header of a TIFF file, making scientific imaging metadata accessible to any software that can read the TIFF format. To support the usage of OME-XML and OME-TIFF, OME has built Bio-Formats, a Java-based library that converts the metadata in proprietary scientific image file formats into the open OME Data Model (*5*).

Bio-Formats is used by many imaging software tools, but because it is written in Java, it has not been included in software that controls most imaging systems. We have therefore ported all the image file input, output and metadata functions in Bio-Formats to C++, in a project called OME Files. OME Files is liberally licensed (BSD) and builds on all major platforms, so it can be used by commercial and academic software projects. Example uses include image

data acquisition, which may require low level access to the acquisition hardware and subsequent writing of the acquired data as OME-TIFF, and image analysis which may require the use of C++ libraries such as OpenCV, Eigen or ITK/VTK, using OME-TIFF as the file format for the input data and any transformed output data. In both of these examples, OME Files will read and write the image data and metadata using OME-XML and OME-TIFF to provide interoperability with the wide range of existing software supporting these open formats. The experimental and acquisition metadata may be stored as OME-XML metadata when writing at acquisition time, and read during subsequent analysis.

OME Files provides a reference implementation in C++ for the OME-XML imaging metadata model and the OME-TIFF image file format (Fig. 1). While OME-TIFF has been adopted by many commercial imaging software providers, since their software is usually written in languages such as C++ or other native programming language, each has implemented its own version of the OME-TIFF readers and writers. OME Files provides a way to minimise this divergence, as it is built and tested on all major compilers and operating systems. In addition, since large-scale calculations most often use native code, OME Files provides a seamless, performant way for these calculations to use OME-TIFF. Finally, OME Files makes it much easier to support OME-TIFF usage in emerging computational environments like iPython notebooks and to explore the development of next generation file formats using containers such as HDF5 to provide scalable access to large and complex image datasets, as well as to support *n*-dimensional datasets (*6*).

Acknowledgements

The authors thank the OME user community for helpful feedback and suggestions for improvements to OME software. This work was supported by a Wellcome Trust Strategic Award (095931/Z/11/Z) and two BBSRC BBR awards (BB/L024233/1 and BB/M018423/1).

Figure Legend

Figure 1. The image acquisition and analysis workflow, focussing on the software and formats that are used. Binary pixel data is combined with experimental and acquisition metadata and, the result stored in either proprietary or open OME-TIFF data files. Software using OME Files (C++) or Bio-Formats (Java) creates OME-TIFF directly. Alternatively, data are written in proprietary or custom formats. OME Files and Bio-Formats, used as libraries inside image processing and analysis software, will read and write OME-TIFF files, ensuring that the processing results are preserved in an open, accessible format. Bio-Formats can

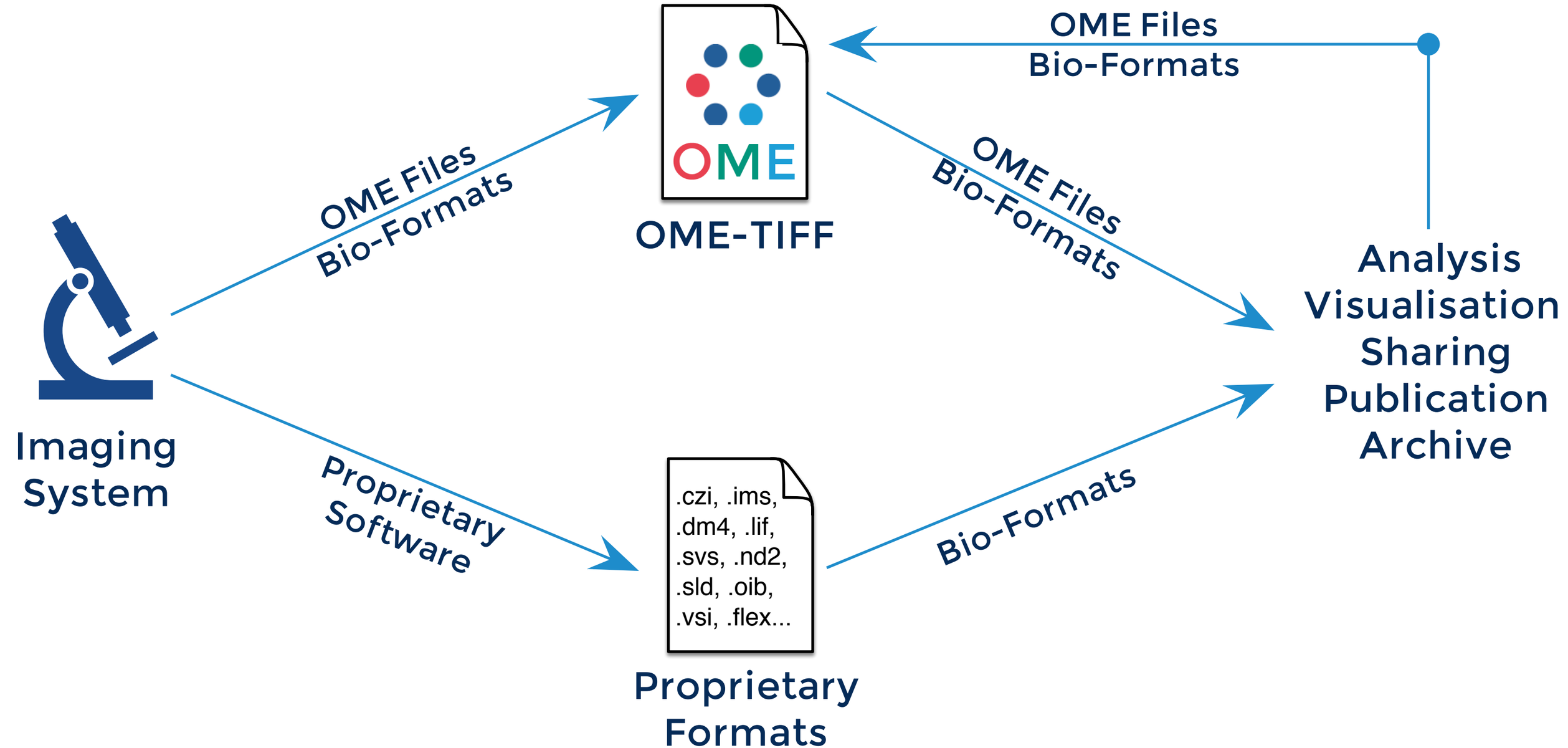
additionally read many proprietary formats, and serves as a way to ensure interoperability between closed formats and other software. It can also be used to convert proprietary formats into open, accessible formats such as OME-TIFF.

Competing Financial Interests

M. L. and J. R. S are affiliated with Glencoe Software, Inc., an open-source US-based commercial company that provides commercial licenses for OME software.

References

1. M. D. Wilkinson *et al.*, The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016).
2. S. Li *et al.*, Metadata management for high content screening in OMERO. *Methods*, (2015).
3. J. R. Swedlow, I. G. Goldberg, K. W. Eliceiri, Bioimage Informatics for Experimental Biology. *Annu Rev Biophys* **38**, 327-346 (2009).
4. I. G. Goldberg *et al.*, The Open Microscopy Environment (OME) Data Model and XML File: Open Tools for Informatics and Quantitative Analysis in Biological Imaging. *Genome Biol.* **6**, R47 (2005).
5. M. Linkert *et al.*, Metadata matters: access to image data in the real world. *J. Cell. Biol.* **189**, 777-782 (2010).
6. T. Pietzsch, S. Saalfeld, S. Preibisch, P. Tomancak, BigDataViewer: visualization and processing for large image data sets. *Nat Methods* **12**, 481-483 (2015).


OME Files
Bio-Formats
OME
OME-TIFF
OME Files
Bio-Formats
OME Files
Bio-Formats
Analysis
Visualisation
Sharing
Publication
Archive
Imaging
System
Proprietary
Software
.czi, .ims,
.dm4, .lif,
.svs, .nd2,
.sld, .oib,
.vsi, .flex...
Proprietary
Formats
Bio-Formats