SPARC Standards for Optical Microscopy Imaging Data and Imaging Metadata

Authors

Susan Tappan, MBF Bioscience; SPARC Data and Resource Center, MAP-Core
Maryann Martone, University of California San Diego; SPARC Data and Resource Center, K-Core

Version 1.0

Approved by the SPARC Data Standards Committee on March 3, 2021

Purpose

The purpose of this standard is to ensure a consistent set of microscopy imaging-specific metadata for primary and derivative imaging data across the SPARC ecosystem is included in experimental datasets. Microscopy imaging data is a common experimental modality, and is inherently diverse. Variability in the experimental sample preparation as well as microscopy settings, methods and equipment all contribute to the complexity. This document establishes a metadata standard so that microscopy data on SPARC.science portal provides relevant context, and imaging metadata standards. Note that this recommendation is microscopy image-specific and is meant to be used in conjunction with the existing SPARC Dataset Minimum Information Standard v3.0 (https://doi.org/10.5281/zenodo.5348190). In addition, only files that are considered to be data products suitable for analysis, not ancillary supporting images for documentation or display are covered by this standard.

Definitions

Derived imaging data: data files that are generated from primary imaging data for the purposes of analysis or FAIR re-use.

Primary imaging data: Data files that are acquired from an optical microscope for the purposes of analysis. These do not include ancillary files created for documentation or visualization purposes, e.g., graphical abstract image.

Background

Essential metadata is considered the critical information necessary to understand and use the image. It includes relevant data about the image itself, as well as how it was acquired. Image information, including the number of pixels in each dimension, file size, and type of image (greyscale, RGB), is required for image display in software applications, whereas sample information is required for the person looking at the image to understand what is being displayed. ***Specifically, without knowing each histologic label or stain captured in each channel of the image as well as the physical dimensionality of the pixel or voxel (in 2D or 3D images respectively), it is simply not possible to understand the content of the image.*** We are used to being provided this sample information when we read figure legends, complete with scale bar, that accompany scientific images (Figure 1); this information must be supplied for a dataset as well. The types of metadata described in this standard ensure that the experimental data generated within the SPARC program have the essential information for use and re-use (Figure 1B).

The proposed standard defines 21 required metadata fields (Table 1) from a total of 47 commonly written metadata fields based on the OME-TIFF specification. Metadata is categorized into 3 types: file metadata, sample metadata, acquisition metadata. Of the 21 required fields, 14 fields typically cannot be edited by users (indicated by Required* in Table 1,

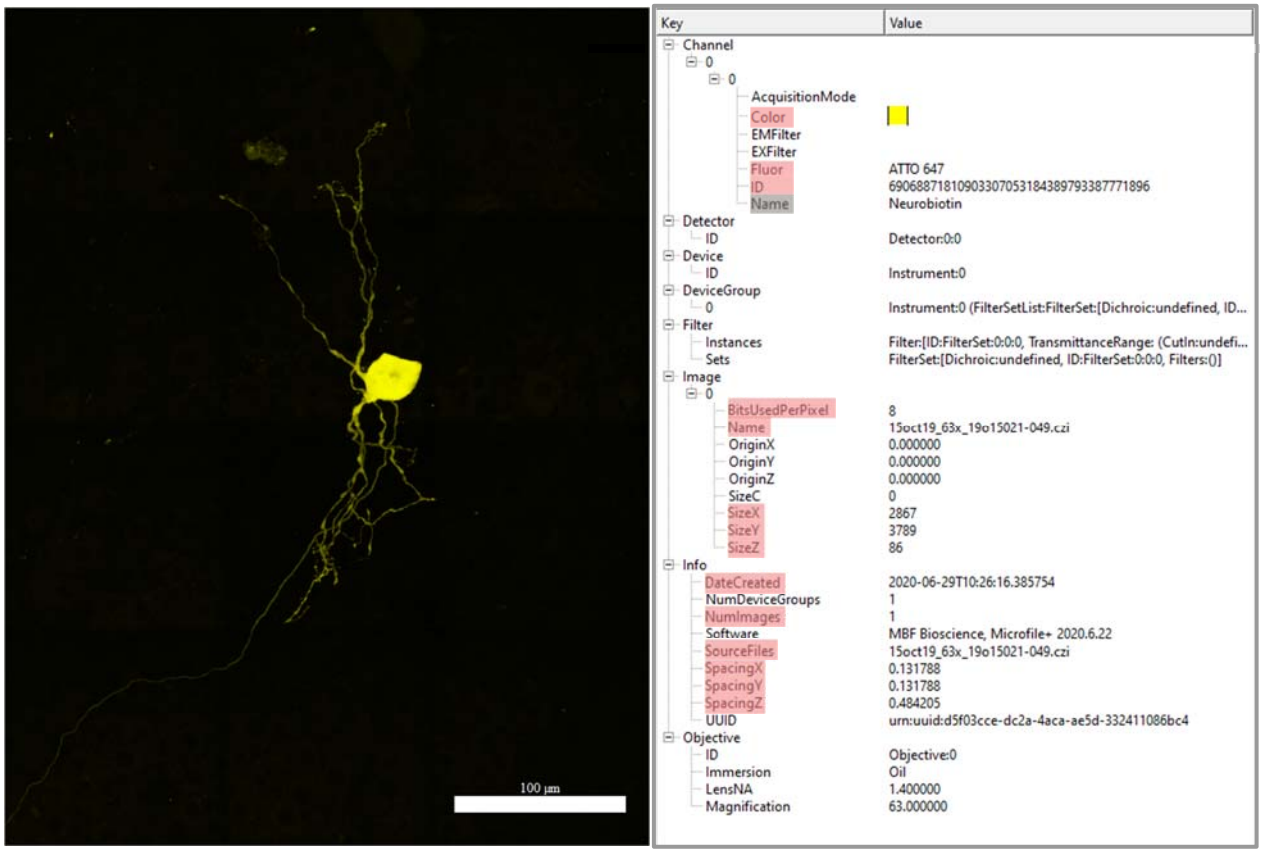

Figure 1 (A) A neuron from the stellate ganglion backfilled with Neurobiotin. A 3D image was acquired on a Zeiss confocal microscope at 63X magnification. Scale Bar equals 100 micrometers. (B) The metadata fields and values for the microscopy image shown in panel A. The required metadata field names are highlighted. The fields with values automatically extracted from the original CZI image are highlighted in red and those manually entered are highlighted in gray. The only manually entered metadata value for this image was the channel name representing the channel target label.

rightmost column), that is, they are automatically output from the imaging device and written to the file header. If so, then only 7 of these metadata elements may need to be entered manually with a typically configured device, and all may be automatically written with an optimally configured device (see next paragraph).

SPARC specification for required and recommended metadata

Metadata Field Name	OME TIFF Field Name	Definition	Example Entry	Present in Source Image	SPARC Metadata Standard
Acquisition Date	DateCreated	Date the new file was saved	2020-06-25T12:38:46.4507	n/a	Required*
Info:SourceFiles	SourceFiles	File extension of source file	.CZI	Automatic	Required*
Info:SourceFiles	SourceFiles	Base file name of source file	VagusPseudoRabiesRat02	Automatic	Required*
Total image size		Size of file (measured in bytes)	4.1 GB	Automatic	Required*
Instrument ID	ID	Microscope manufacturer part number	430037-9021-000	When Configured	Optional
Microscope Manufacturer	Manufacturer	Microscope manufacturer	Carl Zeiss	When Configured	Optional
Imaging Device Model	Model	Microscope model	Axiolab 5 TL+FL	When Configured	Optional
Imaging Device Type	Type	Microscope type	Upright	When Configured	Optional
Software	Creator	Software used to acquire and/or convert file	Zeiss Zen Blue, MBF Bioscience Microfile+ 2020.6.22	When Configured	Optional
PhysicalSizeX	PhysicalSizeX	Size of pixel in X (measured in micrometers)	0.131788	When Configured	Required
PhysicalSizeY	PhysicalSizeY	Size of pixel in Y (measured in micrometers)	0.131788	When Configured	Required
PhysicalSizeZ	PhysicalSizeZ	The physical size of a pixel in the Z direction (µm).	0.484205	When Configured	Required
SizeX	SizeX	Number of Pixels in X	3789	Automatic	Required*
SizeY	SizeY	Number of Pixels in Y	2867	Automatic	Required*
SizeZ	SizeZ	Number of image planes	77	When Configured	Required*
NumImages		For source files that are comprised of many images, identification of how many images	1	Automatic	Required*
SizeC	SizeC	Number of color channels	3	Automatic	Required*
SizeT	SizeT	For source files that are container files, where files are individual time points, identification of how many time points	3	When Configured	Optional*
ZStep	ZStep	The distance (µm) between image planes in the Z direction when different than SizeZ	60		Optional

Objective ID	ID	Objective manufacturer part number	MRD00205	When Configured	Optional
Objective CalibratedMagnification	CalibratedMagnification	Objective magnification	20X	When Configured	Required*
Objective Model	Model	Objective model	CFI Plan Apo Lambda 20X	When Configured	Optional
Objective Manufacturer	Manufacturer	Objective manufacturer	Nikon	When Configured	Optional
LensNA	LensNA	Numerical aperture of objective lens	0.45	When Configured	Optional
Immersion	Immersion	Immersion media used while imaging	Air	When Configured	Optional
ObjectiveSettings RefractiveIndex	RefractiveIndex	Refractive index of immersion media	1	When Configured	Optional
SignificantBits	SignificantBits	Bit depth per channel	8	Automatic	Required*
Detector Name	Model	Camera model	AxioCam	When Configured	Optional
Detector Manufacturer	Manufacturer	Camera manufacturer	Zeiss	When Configured	Optional
Detector ID	ID	Camera part number	412-312	When Configured	Optional
Camera Type	Type	Camera type	Monochrome	When Configured	Optional
PMT:Manufacturer	Manufacturer	Photomultiplier tube manufacturer	Hammamatsu	When Configured	Optional
PMT:Model	Model	Photomultiplier tube model	R550	When Configured	Optional
AcquisitionMode	AcquisitionMode	Imaging modality	Widefield	When Configured	Required*I
Channel ID	ID	Channel Count	0, 1, 2, 3...	Automatic	Required*
Channel Name	Name	Channel target label	Neurobiotin	when Configured	Required
Channel Fluor	Fluor	Channel target fluorophore	ATTO 647	When Configured	Required
Channel Color	Color	Pseudocolor assigned to channel display	yellow	When Configured	Required
Channel EmissionWavelength	EmissionWavelength	Emission filter center wavelength	700	When Configured	Optional
Channel ExcitationWavelength	ExcitationWavelength	Excitation filter center wavelength	647	When Configured	Optional
NDFilter	NDFilter	Neutral density filter's average transmitted light percentage	40	When Configured	Optional
Exposure	ExposureTime	Exposure in seconds (written per channel)	1.5	When Configured	Optional
ContrastMethod	ContrastMethod	The technique used to achieve contrast for each channel	Fluorescence	When Configured	Optional
Info:PostProcessing	Info:PostProcessing	Type of post processing	Deconvolution	When configured	Optional
Info:CompressionType	Compression	Image compression type	JP2000, none	Automatic	Required*

Info:CompressionValue	Value	Image compression ratio (ex. 20:1)	20	Automatic	Required*
Info:Misc	Info:Misc	Label for brightfield or widefield RGB images, where label does not correlate to specific channel.	hematoxylin and eosin		Required
OrigDisc		All existing metadata from the source image is conserved and stored in this section	All original metadata will be preserved by the MicroFile+ converter to retain all information even if it is unable to be extracted to individual fields	Automatic	Required*

Table 1. Definition and example entry of 46 commonly written optical microscopy imaging metadata fields based on the OME-TIFF specification. If the metadata field is present in source imagery and does not require device configuration, the entity is described as “Automatic” in the **Present in Source Image** column. All other entities that are present upon device configuration are described as “When Configured” in the **Present in Source Image** column. Required or optional metadata fields for the SPARC Standards for Optical Microscopy Imaging Data and Imaging Metadata are indicated in the **SPARC Metadata Standard** column.

Note that experimental metadata, including subject ID, species, sex, region, and experimental protocols are covered by other aspects of the SPARC data structure and are not duplicated here. For SPARC, the imaging data must be organized according to subject or sample ID.

Modern image acquisition software often can be configured to write expansive information about the equipment details and settings used as well as information about the sample under view directly to the file at the time of capture. Thus, recommended best practice is to configure the microscope software to include this information, and to use this native format whenever possible for image analyses. Often, when image data is saved in another form, such as going from the Zeiss standard (.CZI) format to an open format such as TIFF, nearly all metadata is lost from the file. When this occurs as a part of the experimental analysis pipeline, a final step to re-create the metadata is required. If this information is not a part of the minimum information available, then the file is ultimately of limited value.

There are at least three ways of addressing the challenges of assisting interoperability with respect to image metadata. The simplest approach is to provide the essential metadata as an independent file (such as a spreadsheet or sidecar). Unfortunately, this increases the likelihood that the metadata may be separated from the image itself since each is contained in a separate file. On the other hand, enveloping the image file and the metadata file within a container file, such as hd5, keeps the information together but it reduces the ease of re-use since image display and analysis software must be able to read the novel format. Some acquisition software, such as Zen Blue, do provide the ability to augment previously acquired images written in Zeiss standard format with missing metadata at a later time. A novel freely available

file converter, MicroFile+, has been created to generate widely supported image formats with the necessary metadata included within the file header when microscopy image files need to be augmented to improve interoperability. By providing a form of the proprietary source file as two open formats inclusive of the essential metadata, OME-TIFF and JPEG2000, we reduce the barrier for re-use.

Required file formats

All microscopy primary imaging data utilized for data analysis purposes will be made available on SPARC.science. All microscopy primary imaging data will be converted to include the required metadata in OME-TIFF and JPEG2000 formats to meet the SPARC Standards for Optical Microscopy Imaging Data and Imaging Metadata. All derived and standardized microscopy imaging data and metadata will also be made available on SPARC.science.

Acknowledgements

We thank the SPARC Data Standards Committee for willing assistance and thoughtful review process.