**Open Reference Data Repository for Nanoscopy**

Advancements in high-resolution microscopy lead to novel imaging methods achieving light-optical resolutions in nanometer scale. This allows scientists to gain new insights on the cell nucleus level. However, due to these high resolutions light-optical microscopy produces tremendous amounts of data that needs to be managed and made accessible in a proper way.

**Open Reference Data Repository**

Novel imaging methods enable new insights but very often the images are hard to interpret and potential users and scientists need to be trained on the basis of reference data and their already known interpretation. These need to be publicly shared in an open reference data repository. Furthermore, new insights of the reference data will be gained with growing experience. Therefore, appropriate tools must enable the open discussion of data and associated analysis results. This necessitates data sharing on the one hand and annotation capabilities on the other. As this discussion may take several years, the data has to be preserved on the content and the bit level for sustainable access.

**Extremely Large Datasets**

The data management of an Open Reference Data Repository is challenging simply because of the size of the datasets. A dataset, covering a single cell, is composed of thousands of high-resolution images summing up to a size of 150-200 TB. Therefore, download and local processing of the data is not feasible and data processing must be enabled by computing resources located near to the repository. As the algorithms used for processing the datasets are also under development detailed provenance information must be captured in order to be able to assess and to retrace obtained results.

**Metadata**

Provenance information is just a small part of metadata needed for the Open Reference Data Archive. Basically, metadata needed by the system can be grouped into two categories: immutable and mutable metadata. The immutable metadata describes e.g. the imaging process, its parameters and the imaged specimen. The mutable metadata covers administrative metadata, e.g. for discovery, retrieval and preservation, as well as the annotation information created by scientists during their discussions about the interpretation of the data stored in the Open Reference Data Respository.