Use Case Description:

The metadata associated with the large datasets in Localization Microscopy (LM) is of vital importance to the researchers. Nanoscopy is a novel imaging technique wherein sub-molecular structures easily viewed via high-resolution images. The generation of the high-resolution images is not a single step process but rather an arduous and work-expensive process. The dataset size for a complete series of measurement can go up to 150-200TB. Therefore it is necessary to correctly identify the metadata which is generated either from the high-resolution microscope or by the researchers. The complete workflow of the LM experiment and the metadata for each step is presented below:

The metadata produced at various stages will be categorized as below [1]:

- Descriptive Metadata – DM
- Structural Metadata – SM
- Administrative Metadata/Technical Metadata – AM or TM

<table>
<thead>
<tr>
<th>Step #</th>
<th>Description</th>
<th>Corresponding Metadata</th>
</tr>
</thead>
</table>
| 1      | Raw-datasets
- The Raw datasets are produced from the high-resolution microscopes.
- For reuse purpose the raw datasets must be archived in large scale data storage using a research data repository system. | 1. [AM] User Authentication metadata (username, password or certificate)
3. [TM] For describing the setup of the high-resolution microscope (for example: Camera setup details )
4. [AM] Experimenter, Organization specific metadata |

2 | Intermediate Results
- The intermediate datasets are generated from the raw-datasets
- Again for reuse purpose the intermediate results are archived in the repository system | 1. [TM] Description of software/algorithm used to produce the intermediate result (for example: algorithm and software details)
2. [AM] Researcher information |
### 3a Generating high-resolution images:
- Producing final high-resolution images from the intermediate datasets
- For sharing and reusing the high-resolution images are archived in the repository system

1. [AM] Researcher’s information that is responsible for producing the high-resolution images from the intermediate results.
2. [TM] Description of software/algorithms used to generate the final high-resolution images.
3. [DM] Metadata describing the high-resolution images and unique-ID for referencing the images.

#### 3b Miscellaneous analysis:
- Various algorithms are executed on the intermediate results to determine specific attributes of the investigation.
- The high-resolution images with the analysis information is generated

1. [AM] Researcher information who was responsible for producing the Analysed images.
2. [TM] Description of software/algorithms used for analysis.
3. [DM] Metadata describing the analysed high-resolution images.
4. **Image Annotation:**
   - Annotation of images for sharing new insights from the high-resolution images (scientific expertise)

5. **Sharing Data (Raw, Intermediate, High-resolution images + annotations)**
   - Raw dataset, intermediate result, high-resolution images including the annotations needs to be shared between the communities and also published in future.

---

1. [AM] Researcher’s information who are annotating the images with scientific content.
2. [DM] Annotating the images with scientific information (for example: vocabulary controlled metadata or free-text)

1. [AM] Researchers/Scientific community’s details. Data sharing metadata (access rights), publication metadata
References