# Microscopy Image Processing Workflow

This workflow provides a complete pipeline for processing microscopy images, including raw data handling, metadata extraction, and image preprocessing. It is designed for high-resolution microscopy research and collaboration.

## Data Management Recommendations

### Raw Data Storage:

* + It is recommended to store the raw data (e.g., CZI files) as a baseline for reproducibility. If raw data is too large (e.g., 500 GB), consider downsampling it for testing and archival purposes.

### Processing:

* + Use the provided processing scripts to perform deskewing, deconvolution, and other preprocessing steps. Note that processed data can become significantly larger (e.g., a 500 GB raw dataset may expand to 700 GB after processing).

### Summarized Data:

* + Instead of storing the full deskewed or deconvolved images, save summarized data to reduce storage requirements. Summarized data could include key metrics, visualizations, or compressed outputs.

## Data & Metadata

Raw data stored in CZI format generated by Zeiss software by researchers/users

* Embedded Metadata: Includes acquisition parameters, sample information, and instrument settings

A screen shot of a computer program

Description automatically generated

* **External Metadata:** Linked CSV files with clinical or experimental sample information. This is generated synthetically by WEHI’s Intern Clinical Dashboard Team

A screenshot of a computer

Description automatically generated

Example how external metadata look like. There are 23 columns information containing patients’ information. This could be extended to other extension such as JSON.

## Image Processing Workflow

1. Deskewing: Corrects image distortions from acquisition.  
2. Deconvolution: Enhances clarity and sharpness.  
3. Downsampling: Reduces resolution for storage efficiency.  
4. Data Summarization: Generates key metrics and compressed outputs.

## Repository Contents

- main.py: Script for processing CZI files into Zarr format.  
- External Metadata File: Links images to sample and patient data.  
- Processed Data: Saved in Zarr/TIFF format for visualization and analysis.  
- Documentation: README file detailing the workflow.

## Data Access

Usage: Download raw or downsampled datasets as needed for analysis.

1. Primary Dataset: Available on Zenodo: <https://zenodo.org/records/7117784>
2. WEHI Internships Dataset: Available on Zenodo: <https://zenodo.org/records/14807924>
3. Data Portal Github: <https://github.com/DBK333/Omero-DataPortal/tree/main/OmeroImageSamples>

## Using the Workflow

- Clone the repository and navigate to the project directory.  
- Run `python main.py` to process CZI files to Zarr format.  
- View processed data with Napari or other visualization tools.  
- Extract metadata using Zeiss software or included scripts.

## Image Preview in Omero

This will only show the preview of images or the summarized images.

A screenshot of a computer

Description automatically generated