# Data Management Plan

## Data Description

Our research project focuses on investigating physical principles of Gastrulation in Tribolium castaneum embryo development using light-sheet microscopy. The primary data consists of 3D time-lapse imaging data stored in the NGFF file format. This data captures embryonic development processes at high spatial and temporal resolution. The data volume is expected to be substantial due to the high-resolution nature of light-sheet microscopy and the temporal extent of developmental studies.

## Documentation and Data Quality

All experimental data undergoes initial quality assessment by two scientists, typically involving a PhD student and either a post-doc or group leader. Quality control procedures are thoroughly documented in an institutional electronic laboratory notebook system. Every step of data acquisition, processing, and analysis is recorded, ensuring reproducibility and transparency. The analysis workflows are documented in Jupyter notebooks, which are stored alongside the code in our version control system.

## Storage and Technical Archiving

Our data storage strategy incorporates multiple secure layers. The primary storage solution is an institutional Omero-Server, specifically designed for microscopy data management. All data is automatically backed up by the university compute center. Analysis is performed through a Jupyter Lab server, with results stored alongside raw data on the Omero-Server. Code is maintained in the institutional git-server with regular backups. Manuscript development occurs in Overleaf with automated synchronization to our git server. This comprehensive system ensures data safety throughout the project's four-year duration.

## Legal Obligations and Conditions

We adhere to clear licensing and data protection protocols. All published data is released under CC-BY 4.0 license, while code is distributed under the BSD3 license. Our institutional laboratory notebook system requires a monthly subscription of 10 euros per user. The university compute center maintains infrastructure supporting these services, with associated costs of 20,000 euros plus half a position for IT support. All data management practices comply with institutional policies requiring minimum 15-year data retention.

## Data Exchange and Long-term Data Accessibility

Our data sharing strategy ensures wide accessibility of research outputs. Upon manuscript submission to bioRxiv, we publish our code on the institutional git-server with appropriate version tagging. Simultaneously, we deposit our imaging data on zenodo.org, ensuring long-term accessibility. Additional copies of all project components - data, code, and manuscript - are preserved on the institutional archive server, guaranteed for 15 years. This multi-pronged approach ensures both immediate accessibility and long-term preservation of our research outputs.

## Responsibilities and Resources

Project responsibilities are clearly distributed among four team members. Data acquisition and analysis are primarily conducted by PhD students and post-docs, with oversight from group leaders. All team members contribute to documentation in the electronic laboratory notebook. The IT department manages infrastructure maintenance and backups, while the compute center ensures long-term storage and server maintenance. Resource allocation includes storage infrastructure costs and dedicated IT support, ensuring sustainable project execution throughout its duration.

---