Data Life Cycle

Description

This page will help users to understand how to manage their data on the cluster. We provide a quick procedure here, but for more details, please consult the Data Management Plan (DMP) provided by Unige.

Each user is responsible for their data and must manage it from generation until deletion from the cluster.

From Terms of use → Storage:

```
The HPC clusters are not a long-term storage provider: users are requested
to manage their files
on a regular basis by deleting unneeded files and migrating results or
valuable data to a permanent location such as Tape NASAC or Yareta.
```

This ensures enough space for everyone and guarantees optimal performance for computing.

Data Management

Below is a schema representing the data life cycle, which includes the following stages:

- **Acquisition:** The process of collecting or generating data.
- **Storage:** The data may be stored on HPC storage for production purposes only (e.g., Home, Scratch, Fast, etc.).
- **Processing:** The manipulation or analysis of data to extract useful information.
- **Usage:** The utilization of processed data for research, analysis, or other purposes.
- **Disposal:** This involves backing up and migrating data to appropriate storage solutions (e.g., NASAC, Yareta), and deleting data from the HPC cluster.

Each stage is crucial for ensuring that data is handled efficiently and responsibly throughout its life cycle.