

## Collection Ingest, Management and Preservation Policy

Long-term viability and access to data and resources is one of the most crucial requirements for research within the digital arts and humanities. Robust policies for collection management and preservation of data will ensure that resources are available for future use by researchers, and will allow discovery and sharing through the DARIAH infrastructure. All users of DARIAH should have an interest in the long-term maintenance of the data they create and use.

### Who is the policy for?

- Partners within DARIAH who are responsible for the collection, management and preservation of data;
- Repositories which are part of the DARIAH network;
- Researchers and DARIAH users looking for guidance.

DARIAH will assist organisations in its network to adhere to best practice in all areas of data management. This will ensure that the processes for ingest and preservation are of the highest possible quality while promoting interoperability across collections. Users will be able to look to DARIAH for guidance on the

procedures that organisations should follow in managing data, so they can make informed decisions when working with repositories.

### DARIAH's collection ingest, management and preservation policy:

- Outlines fundamental processes for ingesting data for use by DARIAH's user community.
- Provides a framework for understanding the processes of the management and preservation of digital data.
- Allows users and organisations to be effective partners in the preservation process.
- Expands knowledge about the necessity of long-term maintenance of data for arts and humanities research.
- Educates researchers and repositories about their roles in data preservation.

Interoperability of data across repositories and collections is a major concern of researchers working in the digital arts and humanities. Proper management of data will ensure that data is held in the necessary formats and has associated metadata to make certain that the data is discoverable and interoperable.

DARIAH's collection ingest, management and preservation policy covers the whole life cycle of digital data and resources, from the pre-ingest phase to the long-term preservation of the data. It aims to ensure that the processes which data managers undergo when presented with research data are robust. It sets out the necessary stages in the life of data to be managed and recommends steps to ensure interoperability.

## Data Management Stages:

- **Pre-ingest:** Pro-active planning of data creation projects, and throughout the data life cycle prior the depositing of data and resources is universally beneficial. It has cost-saving benefits because it ensures the quality and suitability of the data and research advantages in producing high-quality, sustainable resources.
- **Ingest:** The data manager should ensure that data is received from the creator in a format which makes long-term preservation possible, gauge the level of metadata provided by the depositor and work with the creator to ensure that the metadata provided is of sufficient quality. It is also vital that all rights be cleared prior to ingest of data. Once this has been accomplished all activities in ingest should be fully documented, and, if necessary the data transferred to preservation formats.
- **Long-term Management:** It is vital to maintain and provide access to the search information about the data. Custodians of the digital objects have a responsibility to store the data in suitable conditions, ensuring the viability of the original deposited information, and to migrate the data to new formats when necessary.

With procedures such as these in place the use, and re-use of high quality resources and research data will allow arts and humanities researchers the materials they require to generate the best possible outcomes.

[www.dariah.eu](http://www.dariah.eu)