

## RELEASE NOTICE

# NorStore Research Data Archive

**NorStore, the national e-infrastructure for scientific data, is a broad research infrastructure which offers services and facilities to all scientific disciplines that require access to advanced, large-scale or high-end resources for storing, archiving, publishing or searching digital databases and collections. In a response to the considerable big data challenges, NorStore is proud to announce the launch of a new national service for archiving and sharing large-scale research data in Norway. The new service, “The NorStore research data archive”, presently offered as a pilot, is the result of a broad community driven request from users to offer services and tools for long-term archiving and open access publishing of research data.**

UNINETT Sigma AS is the national e-infrastructure coordinator in Norway, charged with seeking funding, user-community interaction, activity planning, implementation of services and international cooperation. Service provision and operation of resources are carried out within a national consortium consisting of the four oldest universities, Univ. of Tromsø, NTNU, Univ. of Bergen and Univ. of Oslo, and coordinator UNINETT Sigma. Funding is provided by the Research Council of Norway (RCN) and the four university partners. NorStore is the national e-infrastructure for large and complex scientific data.

NorStore data management is conceptualised by a data life cycle wheel, representing the various stages of data handling within the context of research and science (see illustration). Thus far, NorStore has offered the community access to resources to facilitate management of data and data processing (e.g. dealing with input or output data from high performance computing). This is known as the ‘project area’ and is now supplemented by a new service – the ‘NorStore research data archive’. This will offer long-term archiving, meta-data provisioning, data sharing, content integrity and curation of legacy research data. In Norway it is a national strategic ambition to provide open access to public data, including research data. NorStore has now developed a framework for handling, documenting and archiving such data. The Archive provides both an interactive and a non-interactive interface to deposit, describe and manage datasets from the user community, accessing the service either through a web-interface or command-lines from the project area.

### **What data is currently stored in NorStore?**

NorStore hosts datasets from a broad variety of research disciplines and a wide spectrum of data types. Data hosted on the e-infrastructure comes from research areas such as climate modeling, sensory data, marine environmental and fisheries data, observations (physics, astronomy, meteorology), medical imaging, numerical simulations, genome sequencing, language technology (e.g. corpus, audio/video recordings) and much more. Common to all are that the data can be characterised as complex or large-scale (repositories, collections) or with a demonstrated need for advanced services to tackle the analysis, processing or preservation of the data.

Published metadata will be provided in an 'open access' framework, while downloading and depositing datasets require Feide/OpenIdP authentication (Feide is the national federated identity service in Norway). In the near future it is anticipated that the archive will provide published datasets with a unique persistent identifier (PID), such that the dataset can be permanently referenced from other sources (e.g. CRISTin). There is currently no national PID service suitable for the use proposed here. Instead that datasets should be published and available for download from the archive for a period of 10 years. Legacy data or larger data collections of national importance may be supported for much longer periods (indefinitely). While waiting for a suitable national PID provider, datasets will not be given an official identifier or published for public access (please see the disclaimer notice in the Terms and Conditions).

Depositing of datasets require user authentication through a supported identity provider (IdP). The Archive currently supports Feide (<https://www.feide.no>) and OpenIdP (<https://openidp.feide.no>). Users without a Feide enabled account can register for an OpenIdP account and select OpenId as a provider. Data can either be uploaded through the online web-interface or deposited from the command line from an existing project area. The Depositor, who authenticates and initiates the dataset, is required to provide meta-data about the origin, discipline, licensing and access rights to the data. It is also a requirement to provide a Data manager and Rights holder for every dataset.

The NorStore Research Data Archive is accessed at the following location:

<http://archive.norstore.no>