

DataverseNO Preservation Policy

 site.uit.no/dataverseno/about/policy-framework/preservation-policy/

Purpose of this Policy

This document describes DataverseNO's commitments and approaches to responsible and sustainable stewardship of published Datasets in the long term. The development of a reliable digital repository that adheres to, and remains compliant with, the standards expounded in the DataverseNO policies remains the best mechanism to support the access and (re)use of research data published in DataverseNO.

The DataverseNO Preservation Policy also addresses a de facto standard of the Digital Preservation community, Trusted Digital Repositories: Attributes and Responsibilities (2002, [PDF](#)). The organization of the policy reflects the seven attributes of a trusted digital repository:

- OAIS compliance
- Administrative responsibility
- Organizational viability
- Financial sustainability
- Technological and procedural suitability
- Systems security
- Procedural accountability

OAIS compliance

In achieving its Digital Preservation objectives, DataverseNO recognises the need to comply with the prevailing standards and practice of the Digital Preservation community. The Reference Model for an Open Archival Information System (OAIS, 2012, [PDF](#)) is an international standard providing common terms and concepts and a framework for entities and relationships between entities in Digital Preservation environments. OAIS is a conceptual framework and not a concrete implementation plan. DataverseNO strives to develop its Digital Preservation policies, repository, and strategies in accordance with the broad guidance given in the OAIS reference model, as outlined in the following description.

Pre-ingest

The pre-ingest function is not explicitly specified in the OAIS model. However, experience has shown that pre-ingest activities and services help ensuring quality, comprehensibility and accessibility of all information packages by enforcing quality assurance and minimum standards at the point of ingest, and thereby reduce costs associated with the ingest phase.

DataverseNO provides data management guidelines, training and consultancy to groups and individuals within the Designated Community about issues such as metadata standards, file formats, and legal and ethical issues. Considering and addressing such issues before data deposit starts may also have an impact on preservation activities.

The [DataverseNO Accession Policy](#) explains what DataverseNO accepts for publication.

The [DataverseNO Deposit Guidelines](#) require documentation as well as preferred file formats for Datasets to be published in DataverseNO. If data cannot be stored in a preferred file format, they can still be published in their original format, but in that case, DataverseNO does not commit to preserve the data in the long term. The preferred file format list is based on recommendations from the Library of Congress as well as other data repositories with significant holdings such as UK Data Archive and DANS.

Ingest

Ingest is the first functional component of the OAIS reference model. It includes the receipt of information from a Depositor and the validation that the information supplied is uncorrupted and complete. This process also identifies the specific properties of the information which is to be preserved; it authenticates that the information is what it purports to be. The first supplied version is known within DataverseNO as the draft version of a Dataset that has been submitted for review/curation (short: Dataset draft). A Dataset draft consist of a metadata record stored in the repository system, a mandatory ReadMe file (documentation), and one or several data files. Each Dataset draft including all its files are assigned each their Digital Object Identifier (DOI). Dataset drafts are reviewed by Research Data Service staff to ensure compliance with the DataverseNO Deposit Guidelines. In the case of non-compliance, the Depositor is asked to make necessary changes and/or additions. Any changes and/or additions will change the first draft version without the original version being kept in the repository.

A Dataset draft has a close correspondence to the OAIS Submission Information Package (SIP). However, the first supplied version of a Dataset is not kept in the case of curation-induced changes. DataverseNO also deviates from the OAIS model by not creating separate Archival Information Packages (AIPs) for storage. Rather, in the ingest phase of DataverseNO, Dataset drafts are prepared as Dissemination Information Packages (DIPs). As argued by other digital archive providers, e.g. UK Data Archive, the construction of a DIP during the ingest process (rather than automatically from an AIP on demand) has considerable benefits for the preservation process. This allows the archive to reduce errors in co-operation with the producer and maximise data usability. It is known that the production of multiple DIPs which are based on different software packages may lead to a loss of integrity in the underlying data.

Archival Storage

The second functional component of OAIS is Archival Storage. Archival Storage manages the digital objects which are entrusted to the archive. In essence, the purpose of archival storage is to ensure that what is passed to it from the ingest process remains identical and accessible. In the OAIS model, this function receives AIPs and DIPs from the ingest function and adds them to the permanent storage facility, oversees the management of this storage, including media refreshment and monitoring. This function is also responsible for ensuring that AIPs can be retrieved.

As mentioned in the previous section, DataverseNO does not follow the OAIS model fully. In the model, there is a 'Provide Data Function' whereby the Access function (see below) of the archive can request AIP transfer from the storage area. In the reference model, this process ensures that end users receive an authentic version of the data collection. However, due to a combination of factors relating to information security, access conditions and usability, preparing DIPs is part of the ingest process in DataverseNO.

Data Management

Data Management is the third major function of the OAIS reference model. It works in conjunction with the archival storage function. It maintains databases of descriptive metadata; supports external finding aids, and manages administrative metadata which support internal operations, including change control.

Version control/change procedures

Ensuring that any alteration to the preserved version of any part of a Dataset is accurately documented is integral to the authenticity of any data collection. DataverseNO distinguishes between two forms of alteration of published Datasets (DIPs):

- minor version change (Definition: when there is a change to the metadata preserved in the repository system);
- major version change (Definition: when there is change to data file(s) or documentation file(s)).

Any change results in a new version of the published Dataset (DIP). Minor version changes are reflected by decimal number changes in the version number included in the Dataset references, e.g. from version 1.0 to version 1.1. Major changes are reflected by integer number changes in the version number, e.g. from version 1.1 to version 2.0.

Data Deaccessioning

In case of compelling reasons, DataverseNO may remove public access to (Deaccession) Data Files and/or Metadata Files in a Dataset from its holdings. Details regarding the process for Deaccession can be found in the [DataverseNO Deaccession Procedure](#).

Access

The Access function of the OAIS model contains the services and functions that make the archival collection and related services visible to end users. End users interact with the archive to find, request and receive Datasets. By default these processes are web-based, but with support by Research Data Service staff.

Apart from the processes that support these three activities (i.e. find, request and receive Datasets), the access function also implements the security that is related to access.

As a main prerequisite for findability, Datasets published in DataverseNO are furnished with standard recovery metadata in accordance with the requirements of DataCite. In addition, DataverseNO commits to facilitating maximum access and use of research data published in DataverseNO in several ways, as outlined in the [DataverseNO Access and Use Policy](#).

Administration

In the OAIS model the administration function manages the day-to-day operations of the repository. In DataverseNO, the roles of this function are distributed across different, clearly defined internal sections at the owner institution and the partner institutions of the repository. The different roles, responsibilities, and tasks relevant for the operation and development of DataverseNO are outlined in the section *Organization of DataverseNO* on the [About](#) page of the DataverseNO information webpage.

Preservation Planning

Preservation Planning is the function responsible for monitoring the OAIS environment and providing recommendations to the repository (through the Administration function) to ensure that materials are accessible to the Designated Community in the long term. This function detects critical information about shifts in the knowledge base of the designated community, allowing the repository to accommodate shifts in technology. In the OAIS model, Preservation Planning includes:

- Monitoring the designated community
- Monitoring technology
- Monitoring the significant characteristics of the repository's contents
- Developing preservation strategies and standards for continuing access.
- Developing packaging designs and migration or routine transfer plans.

The approaches and commitments of DataverseNO to preservation planning for its digital contents are described in this document (DataverseNO Preservation Policy).

Administrative Responsibility

The second attribute of a trusted digital repository is administrative responsibility. DataverseNO was established in 2017 by UiT The Arctic University of Norway as a national generic repository for open research data. The DataverseNO repository mission complements the broader teaching, research, and public service mandate of its owner, UiT The Arctic University of Norway, as well as their commitment to Open Science. In their national policy for research data management in Norway (in Norwegian, [PDF](#)), the Norwegian Ministry of Education and Research mandate the higher education and research sector in Norway to increased reuse of research data. In the policy, DataverseNO is mentioned as one out of five national, generic research data services to support this mandate.

Mandate and Commitments

Based on its steering documents, DataverseNO upholds the following commitments:

- **Commitment to teaching and research:** The DataverseNO repository supports and enhances teaching and research by providing long-term access to digital research data assets according to national and international standard and best practice recommendations.
- **Open Access practices and advocacy:** By default, the DataverseNO repository makes digital research data assets freely and publicly accessible to the scholarly community as well as the greater public. It also offers access to reliable archival infrastructure to encourage researchers to provide open access to their data to the benefit of the Designated Community of repository as well as to the scholarly community as a whole.
- **Membership services:** DataverseNO preserves digital research data assets and provides its partner institutions with long-term access to their digital collections within the repository.

Preservation Objectives

DataverseNO commits to facilitate that published data remain both authentic and accessible in a long-term perspective. In order for digital records to be described as authentic over time they must adhere to three essential characteristics: be reliable, have integrity and be usable.

Reliability is ensured through the operation of transparent and fully documented preservation strategies, and the provision of metadata required to describe the content, context and provenance of the record.

Integrity is ensured through bit-level preservation and through the provision of metadata to describe all authorised actions undertaken in the course of content and bit-level preservation.

Usability is ensured through content preservation, and the provision of metadata sufficient to allow the record to be located, retrieved and interpreted.

To achieve its preservation objectives, DataverseNO shall:

- Provide and participate in a reliable preservation environment for published Datasets in collaboration with other national and international stakeholders.
- Provide reliable and consistent access to published Datasets available through the repository to Depositors, the Designated Community, the larger research community, and the greater public according to the DataverseNO policies.
- Adapt preservation strategies to incorporate the capabilities afforded by new and emerging technologies in cost-effective and responsible ways.
- Serve the needs of the Designated Community and the partner institutions by enabling uninterrupted access to digital research data assets over time as the technology for digital content creation and distribution evolves.
- Meet the archival requirements of funding agencies committing to the long-term preservation of digital research data assets.
- Demonstrate auditable compliance with and contribute to the development of the standards and practice of the Digital Preservation community.
- Foster collaborative partnerships with the different user groups of the Designated Community and other digital archives to make the best use of available resources and avoid duplicative efforts.

Constraints

In general, to achieve the preservation objectives outline above, preservation decisions at the DataverseNO repository must always be made within the context of its Accession Policy, balancing the constraints of cost alongside the requirements of preservation levels described in the present Preservation Policy.

In particular, in cases requiring compromise, DataverseNO prioritizes the preservation of the content, context, and structure of its digital records, as opposed to preserving their appearance or behaviour.

Organizational Viability

DataverseNO accepts, preserves, manages, and distributes digital research data assets in accordance with the DataverseNO Policy Framework. The Digital Preservation processes and procedures employed by DataverseNO and described in this Preservation Policy demonstrate an explicit institutional commitment to the long-term preservation of and access to its digital data holdings now and into the future.

Scope

The DataverseNO repository accepts responsibility for preserving and making available digital research data, associated documentation, and other metadata provided to the repository by Depositors in accordance with the [DataverseNO Accession Policy](#). Furthermore, DataverseNO makes in this Preservation Policy an explicit statement on what levels of preservation apply to specific types of digital objects.

Operating Principles

The DataverseNO repository builds on an established set of principles upon which its Digital Preservation program has been developed and implemented. DataverseNO strives to:

- Endeavour to model services in accordance to the National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation ([PDF](#)) and the Reference Model for an Open Archival Information System (OAIS, 2012, [PDF](#)) for the long-term preservation and access of digital collections whenever possible.
- Comply with emerging international standards for managing a reliable digital repository infrastructure.
- Adhere to prevailing Designated Community standards for preserving access to Datasets whenever possible.
- Enforce data and metadata quality standards to sustain and enhance the value of data, and facilitate discovery, access, interoperation, and reuse of the data.
- Participate in the development and implementation of standards.
- Commit to an interoperable, scalable digital repository with appropriate storage management for Datasets.
- Create policies, procedures, and practices that are clearly documented and consistent.
- Regularly review and revise policies, procedures, and practices as needed.
- Maintain hardware, software, and storage media containing archival content in keeping with prevailing best practices.
- Establish procedures that meet archival requirements, such as clearly documenting and maintaining provenance, chain of custody, authenticity, and integrity of Datasets.
- Comply with copyright, intellectual property, and ownerships rights and/or other legal rights related to copying, storage, modification and use of digital resources.

Roles and Responsibilities

The different roles, responsibilities, and tasks relevant for the operation and development of DataverseNO are outlined in the section *Organization of DataverseNO* on the [About](#) page of the DataverseNO information webpage. As the owner of the DataverseNO repository, UiT The Arctic University of Norway has accepted the overall responsibility for the Digital Preservation program of the repository.

The Digital Preservation program of DataverseNO is carried out as a distributed function that is integrated into operations across the organization of DataverseNO. DataverseNO has identified the following stakeholder categories involved in the Digital Preservation program:

- **Depositor:** Person(s) who provide the information to be preserved. Depositors are members of the Designated Community of DataverseNO. Depositors are responsible for complying with established deposit requirements and working with the Research Data Service staff of the repository to ensure a successful data deposit.
- **Curator:** Research Data Service staff taking care of ongoing curation of specific collections. Curators check deposited Datasets for compliance with the DataverseNO policies and guidelines, and provide guidance to Depositors on how to adjust deposited Dataset to become compliant with these policies and guidelines before the Datasets are published by the responsible curator. Curators also take care of specific long-term preservation operations as specified by the repository management and the collection management.
- **Collection Management:** Research Data Service staff taking care of the management and operation of their collection. The collection management are responsible for specific long-term preservation operations as described in this Preservation Policy, and further specified by the repository management.
- **Repository Management:** Research Data Service staff employed at the owner institution of DataverseNO taking care of the management and operation of the DataverseNO repository. The repository management take care of the establishment, review, revision, and implementation of the DataverseNO preservation policy, including the long-term preservation operations not delegated to the collection management.
- **Advisory Committee:** The advisory committee for DataverseNO, and the advisory committees for collections within DataverseNO give advice to the repository and collection management as well as to the Board of DataverseNO on any aspects of Digital Preservation relevant for the repository.
- **Board:** The Board of DataverseNO have the overall responsibility for all aspects of the DataverseNO preservation policy, and for developing and keeping DataverseNO abreast of the challenges of Digital Preservation in a long-term perspective.

Selection and Accession

The DataverseNO repository identifies and solicits contributions of research data considered by the Designated Community to be of significance to their fields of study. Data to be published in DataverseNO are accepted in accordance with the appraisal criteria as set forth in the [DataverseNO Accession Policy](#). The [DataverseNO Deposit Guidelines](#) provide guidance for Depositors to encourage complete and well-documented deposits. These guidelines also include instructions on how to deposit research data that cannot be made available immediately (cf. Embargo).

Access and Use

The DataverseNO repository defines its Designated Community as consisting primarily of the members of the user groups falling into the following three main categories:

1. researchers from Norwegian research institutions that are partners of DataverseNO
2. researchers working within the scope of any special collection within the DataverseNO repository
3. researchers from Norwegian research institutions that are not partner of DataverseNO

These members include research faculty, students, and other individuals that participate in academic research. Since the DataverseNO repository provides free and open access to its collections, data are not only accessed by the primary user groups described above, but also by other stakeholders in society reliant on access to knowledge, e.g. journalists, teachers, industry as well as the greater public.

As a main prerequisite for access and use, Datasets published in DataverseNO are furnished with standard recovery metadata (including terms of use) in accordance with the requirements of DataCite. In addition, DataverseNO commits to facilitating maximum access and use of research data published in DataverseNO in several ways, as outlined in the [DataverseNO Access and Use Policy](#).

Challenges and Risks

DataverseNO acknowledges the challenges and risks to long-term Digital Preservation. Despite these challenges and risks, DataverseNO commits to provide long-term access to digital data assets to support the research community. The DataverseNO Digital Preservation program addresses these challenges and risks as outlined below:

- **Changes in technology:** Like any organizations engaged in Digital Preservation, DataverseNO needs to be responsive to continually changing technology. As information technology evolves, new digital content types, new capabilities, and new preservation challenges emerge and existing digital content faces the risk of obsolescence. Therefore, DataverseNO continually monitors and responds to changes in technology.
- **Shifts in normative research practice:** As the adoption of novel research techniques become widespread, DataverseNO must become aware of and understand the new tools, practices, and data types these novel techniques yield. DataverseNO adapts its preservation program to accommodate these and other shifts in scholarly practice.

- **Training and awareness:** Most of the Research Data Service staff involved in DataverseNO contribute directly and indirectly to the Digital Preservation program of the repository. DataverseNO is committed to providing appropriate training for and raising awareness about Digital Preservation issues and developments both for Research Data Service staff involved in the operation of the repository, as well as for the Designated Community of the repository.
- **Expansion of roles and responsibilities:** The role of the repository is as dynamic as the landscape in which it serves. Changes in technology, research practices, domain definitions, and stakeholder expectations require Research Data Service staff involved in DataverseNO to receive appropriate training and professional development opportunities to be able to expand roles and responsibilities in order to effectively develop, implement, and maintain a comprehensive Digital Preservation program.

Financial and Organizational Sustainability

DataverseNO is organized in a way that ensures sufficient funding for the operation and further development of the repository in a long-term perspective.

General Context

Both the owner institution and the partner institutions of DataverseNO are state-owned universities and thus part of the national, governmental higher education and research system and under the ultimate responsibility of the Norwegian Ministry of Education and Research. They are all reputable institutions that have existed for many decades – though in some cases not under their current name. Thus, they all are organized and funded in a way that ensures the operation of sustainable services for higher education and research in a long-term perspective. Also, all institutions involved in DataverseNO have recognized Open Science as an important issue in their missions.

As is the case for any other sustainable service – both the owner institution and the partner institutions of DataverseNO allocate their funding and resources to the operation and development of DataverseNO on a scalable basis, but always to a sufficient extent in order to completely fulfil their commitments at any time. This means, e.g., that a partner institution does not allocate all their research support staff to the operation of their institutional collection within DataverseNO right from the establishment of the collection. Allocation of resources on a scalable basis means that necessary funding and staff are allocated gradually as data deposit into the collection increases. This scalable model has proved to be very successful and sustainable in the development and operation of similar services at higher education and research institutions in Norway.

Furthermore, although the resources needed e.g. for data curation increase as more researchers at DataverseNO partner institutions choose to deposit their data into DataverseNO, DataverseNO expects, and has already experienced, that the average time used

on data curation per Dataset will decrease as researchers become more proficient in research data management the more Datasets they have deposited into the repository and the more research data management training they have received at the partner institution or elsewhere.

Institutional Commitment

Owner of DataverseNO

UiT The Arctic University (owner of DataverseNO) has a long-standing record as a pioneer in promoting Open Access, Open Data and Open Science, and has as a goal in its present strategy (2018-2022) to be nationally leading in Open Science ([PDF](#)). Thus, there is a strong commitment at the institution to long-term support, strategic priority and sustainable funding of activities and services like DataverseNO, for the benefit of the institution. In particular, UiT commits to the partner institutions and the Designated Community of DataverseNO to ensure the proper management and operation of DataverseNO in a long-term perspective, and in accordance with the responsibilities described in the [Steering Document for DataverseNO](#).

Partner Institutions

By signing the partner agreement, the partner institutions of DataverseNO commit to operate their institutional collections according to DataverseNO policies and guidelines. This implies that they have to ensure sufficient funding and resources as well as sufficiently qualified staff to fulfil these requirements at any time.

Funding model of DataverseNO

UiT The Arctic University of Norway as the founder and owner of DataverseNO is responsible for the basic funding of the repository. The partner membership fees cover UiT's overhead expenses for offering DataverseNO to their partner institutions. These overhead expenses are related to the management, the operation, and the development of the repository, but not to data curation of any sort – since data curation is the responsibility of the partner institutions. Each partner institution covers their expenses for competence building and attending meetings.

Cooperation and Collaboration

DataverseNO acknowledges Digital Preservation as a shared community responsibility. UiT The Arctic University of Norway (owner of DataverseNO) has long-standing and emerging partnerships with similarly committed organizations (e.g. the National Library of Norway and the Norwegian Centre for Research Data) and is committed to collaborating with other organizations and networks to:

- advance the development of the Digital Preservation program
- share lessons learned with other Digital Preservation programs
- extend the breadth of our available expertise

Generally, in working, cooperating and collaborating with others, DataverseNO desires to:

- understand the goals, objectives, and needs of the communities of creators and the communities of consumers of its digital resources
- identify appropriate partners and stakeholders to contribute to national and international efforts in Digital Preservation
- help develop national and international strategies and initiatives that enable the distribution of collecting, description, service delivery, digitization and preservation activity
- work actively with creators of digital materials to encourage and promote standards and practices

Technological and Procedural Suitability

The DataverseNO repository employs several Digital Preservation strategies and techniques to achieve its preservation objectives. These strategies and techniques are informed by guidelines and procedures used by the Digital Preservation community and developed to align with established archival standards and best practices.

The content of the DataverseNO repository collections consists of digital research data from different academic disciplines. These data come in different data types which are represented in different file formats. The documentation included in each Dataset contains information required to identify, verify, interpret, and use the data.

Deposit Requirements

According to the [DataverseNO Accession Policy](#), the [DataverseNO Deposit Agreement](#), and the [DataverseNO Deposit Guidelines](#), Datasets to be published in DataverseNO must fulfil a number of requirements to support long-term preservation of the digital materials, including the following:

- Each Dataset must include metadata and a ReadMe file containing information required to identify, verify, interpret, and use the data.
- Data Files have to be in preferred file formats suited for long-term preservation as advised on by the repository.
- The Depositor grants DataverseNO the right to convert the deposited Data Files and/or Metadata Files to any medium or format and make multiple copies of the deposited Dataset for the purposes of security, back-up, and preservation.
- For the same or other purposes, the Depositor grants DataverseNO the right to make changes to Descriptive Metadata.

DataverseNO provides information about preferred file formats through a [list](#) published in the DataverseNO Deposit Guidelines as well as through advice during data curation. The provided list and advice are based on the following criteria:

- Openness of the format: Is the format well described and is documentation available? Is the format subject to any patents? Is a license or permission required to use the format?
- Distribution of the format: Is the format used widespread? Will many software programmes be able to understand the format?
- Acceptance of the format as a preservation format: How is the format evaluated on corresponding lists of recommended formats?
- Dependency of the format on external sources of information, for example fonts or pictures with external references.
- If applicable, does the format use standard character encoding?

Preservation Strategies

The deposit requirements described above form the foundation of the preservation program of DataverseNO. Based on these requirements, DataverseNO employs several Digital Preservation strategies that require that all reasonable efforts be applied to ensure the integrity, authenticity, and completeness of the digital content published in the repository. DataverseNO prioritizes the preservation of the content, context, and structure of the published materials, as opposed to preserving their appearance or behaviour.

DataverseNO utilizes the following preservation strategies:

Normalization: On deposit, Research Data Service staff responsible for the curation of the DataverseNO collection in question check Data Files and metadata for completeness and integrity and, if necessary, communicate with Depositors to request changes and/or additions to the deposited Dataset in order to ensure compliance with the deposit requirements described above. This may include the provision of a missing ReadMe file and/or the data in preferred file format(s).

This type of normalization effort may also take place after publication as part of the long-term preservation activities carried out by DataverseNO. In particular, this is the case when normalization of file formats for some reason was not possible or feasible at the time of publication. When carried out after publication, changes made due to normalization result in a new version of the Dataset. The metadata of the Dataset are updated to indicate the reason for any necessary changes that may have been made to the Dataset in the process of normalization.

Format Migration: When DataverseNO identifies part(s) of its content being stored in format(s) that are at risk of obsolescence, a new version of this content will be created in a format more suited to long-term preservation and use. This transformation may consist of

migration to a newer version of the content's existing format, or transformation to a different format altogether. In all cases, preservation of the object's content, context, and structure will be prioritized over the preservation of a specific presentation style or behaviour.

Migrated files are stored alongside the original files and form a new version. The technical metadata of the object are updated to indicate the reason for the change and any necessary changes that may have been made to the file in the process of migration.

The repository management provides a report for the collection managements containing an overview of files of each format within the collection so that objects in need of migration can easily be identified.

Bit Stream Copying: DataverseNO maintains regularly scheduled backups of all information contained in the repository, for use in the event of data loss. In combination with regular fixity checks, which identify potentially damaged content, this process ensures the integrity of content in the DataverseNO repository, and provides a foundation for disaster recovery.

Fixity Checking: All materials in the repository are subject to regular fixity checks, i.e. comparisons of checksum values calculated at a given point in time with those generated at the materials' time of ingest. This activity, when combined with bit stream copying, mitigates the risk of objects becoming corrupt in the repository, as it enables the repository management to identify damaged or corrupted content, and to revert to a valid version of the object from a previous point in time.

Levels of Preservation

The preservation strategies described above are applied to digital objects in the repository at three preservation levels according to the type of file format these objects are represented in. This section describes the preservations levels, the access goals for each object group, and the success measures for each access goal.

Access goals describe the type of activity that users of the preserved content are anticipated to be able to perform with that content in the future.

Success measures identify activities that DataverseNO Research Data Service staff or external organizations can perform on a routine basis to verify that access goals can be achieved.

Preservation Level 1:

- **Object Group:** All objects.
- **Applied preservation strategies:** Bit Stream Copying, Fixity Checking.
- **Access Goals:** Authorized users can access copies of the object in the same format it was originally in the last published version. Preservation at level 1 does not ensure that files are accessible in the same software used at time of access.

- **Success Measures:** Checksum at time of original processing is the same as at time of future access.

Preservation Level 2:

- **Object Group:** All objects.
- **Applied preservation strategies:** Normalization.
- **Access Goals:** Authorized users can get a copy of the data and documentation files that make up a Dataset in a preferred file format that was current at time of capture or ingest, with significant characteristics of the original as represented in the last published version reasonably intact.
- **Success Measures:** The normalized versions of all files that make up a Dataset have checksums that are identical to the ones derived at the time of normalization.

Preservation Level 3:

- **Object Group:** Objects in preferred file format(s).
- **Applied preservation strategies:** Format Migration.
- **Access Goals:** Authorized users can access the resource in file formats that are current at the time of access. Files may not correspond one-to-one with the original files, but the significant characteristics of the original resource as represented in the last published version will be reasonably intact.
- **Success Measures:** The migrated version of the resource retains as many of the significant characteristics of the obsolete version as is practical. Migrated versions of the original are usable in software common at time of access. Migrated versions of all files have future checksums that are identical to the ones derived at the time of migration.

Format Migration is similar to the part of Normalization applying to file formats. However, Format Migration is only offered as a preservation strategy for normalized objects. If an object for some reason has not been normalized and the employed file format becomes obsolete, DataverseNO does not commit to further preservation efforts for the object other than the efforts provided at Preservation Level 1.

Significant Characteristics

Significant characteristics are characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record (Wilson, 2008, p. 15).

To the extent possible given the constraints defined in the present Preservation Policy, DataverseNO attempts to preserve the significant characteristics of the objects contained in published Datasets. As a main rule, DataverseNO employs the significant characteristics as defined by Archivematica ([link](#)). These represent a well-researched, community-led,

operationally tenable set of characteristics. In cases requiring compromise, transformations that maintain the content of the object will be prioritized over those that preserve the presentation and behaviour of the object. In case of uncertainty about the significant characteristics of an object to be preserved, DataverseNO strives by reasonable efforts to obtain advice in the matter from the Depositor or from other representatives of the user group at stake.

Deaccessioning

In case of compelling reasons, DataverseNO may remove public access to (Deaccession) Data Files and/or Metadata Files in a Dataset from its holdings. Details regarding the process for Deaccession can be found in the [DataverseNO Deaccession Procedure](#). The DataverseNO Preservation Policy does not apply to deaccessioned files.

Continuity of Access

In the unlikely case that DataverseNO is closed down UiT The Arctic University of Norway (owner of DataverseNO) commits to ensure that published data are retained and transferred to (an) approved repository/-ies before the service is discontinued. Datasets in institutional collections are transferred to (a) certified general research data repository/-ies. Datasets in special collections are transferred to certified subject-based repositories after consulting the involved user groups. In addition, and according to Norwegian legislation, research data from the governmental sector will be transferred to the [National Archives of Norway](#) for securing long-term availability and accessibility of the data.

Planning and Monitoring

The preservation program described in this preservation policy is implemented in the [DataverseNO Preservation Plan](#). This regularly updated document describes actionable steps to be taken to preserve published Datasets within the DataverseNO repository in the long term.

In order to keep its approaches to preservation and the actionable preservation activities up to date, DataverseNO monitors the development of relevant aspects of its preservation policy, including:

Community: DataverseNO monitors the development of its Designated Community through substantial contacts, for instance through pilot studies with data producers, data curation, training and consultancy, participation in national and international research infrastructures and networks, and by offering domain-specific services.

Technology and procedures: DataverseNO monitors the development of national and international standards, recommendations, procedures and tools for Digital Preservation by following and participating in relevant communities and networks. In particular,

DataverseNO monitors developments in the area of file formats to determine if/when formats preserved within the repository are in need of preservation actions. This monitoring takes place through engagement in the Digital Preservation community, and by using the [Library of Congress's Sustainability of File Formats pages](#), the UK National Archive's [PRONOM](#) service, but also through monitoring of mailing lists and journals related to Digital Preservation.

Repository: DataverseNO monitors the development of the DataverseNO repository to check whether its technical infrastructure and content comply with the chosen requirements for long-term preservation.

Monitoring is carried out at all levels within the DataverseNO organization, including curators, collection managers, and repository managers. The roles and responsibilities for actionable preservation steps are described in the DataverseNO Preservation Plan.

Communication and Training

The preservation program described in this preservation policy including the preservation plan are communicated to Research Data staff at DataverseNO partner institutions through the [DataverseNO Curator Guidelines](#). New Research Data Service staff members are trained in these guidelines within a few weeks after joining the Research Data Service team at a DataverseNO partner institution. Those parts of the preservation measures concerning the overall management of the DataverseNO repository are communicated to the collection managers of DataverseNO through the [DataverseNO Administrator Guidelines](#) as well as through direct contact by the repository management.

Systems Security and Disaster Recovery

UiT The Arctic University of Norway (owner of DataverseNO) is committed to sustaining an effective Digital Preservation infrastructure for its digital collections, which includes the adequate provision of appropriate technologies.

Datasets deposited in DataverseNO utilize the centralized back-end storage and management services at UiT. This is a common storage and management infrastructure for digital collections of enduring value to UiT, covering digitized and born-digital books, manuscripts, photographs, audio-visual materials, scholarly publications, and research data.

DataverseNO is running on UiT's centralized storage and virtualization infrastructure which also hosts the accounting and payroll systems for the whole institution. Any content in DataverseNO is backed-up using an enterprise class backup system with retention policies ensuring that multiple copies are maintained of all data in the system. Data recovery is available from backup as necessary as for other storage services for researchers and students at UiT. The underlying hardware is mirrored between two datacentres in separate buildings

on the UiT campus, where data is replicated to avoid data loss in case of physical threats like fires, floods etc. Both datacentres are secured with at least two layers of key access doors from public areas, and access is restricted to authorised operational staff.

All systems (included DataverseNO) and services delivered by UiT are subject to risk and vulnerability analysis at implementation, at start up, and at regular intervals throughout the lifetime of the systems and services.

The infrastructure and services are revised regularly according to the UiT quality control system. The storage systems are renewed every 6-8 years, which minimizes the risk for long-term deterioration of storage media. The transfer of data from old to new storage systems includes checks for bit-correctness of all data. DataverseNO complies with the UiT requirements for good computer use practices. UiT has developed extensive technical and administrative procedures to ensure consistent and systematic information security. Good practice requirements include system security requirements, operational requirements and regular auditing and review.

Procedural Accountability

DataverseNO is dedicated to promoting trust with its Designated Community, collaborative partners, the professional repository community, and the larger scholarly community through self-assessment, audit, and transparency of policies and procedures.

Audit and Transparency

DataverseNO strives to follow best practices for digital repositories as specified in the de facto standard of the Digital Preservation community, Trusted Digital Repositories: Attributes and Responsibilities (2002, [PDF](#)). As part of these efforts, DataverseNO commits to self-assessment and audit. To demonstrate this commitment, the DataverseNO repository earned the [CoreTrustSeal](#) for compliance with international requirements for sustainable and trustworthy digital data repositories.

DataverseNO routinely revisits and adjusts its policies and procedures to remain responsive to changes and advances in accepted Digital Preservation standards and best practices.

DataverseNO makes its policies, procedures, and results of assessments, including audit trails publicly available online via its information website (<https://info.dataverse.no/>).

Policy Administration

Being part of the DataverseNO Policy Framework, the Dataverse Preservation Policy is subject to three-year review or upon the emergence of new needs, standards and best practices, whichever may come first, and according to the guidelines in the DataverseNO Policy Framework.

Acknowledgements and References

The references below point to documents and resources which the present document is adapted from and inspired by, or which are otherwise referred to in the present document.

Archivemata Significant characteristics.

https://wiki.archivemata.org/Significant_characteristics

Becker, C., Kulovits, H., Guttentbrunner, M., Strodl, S., Rauber, A., & Hofman, H. (2009). Systematic planning for Digital Preservation: evaluating potential strategies and building preservation plans. *International Journal on Digital Libraries*, 10(4), 133–157.

<https://doi.org/10.1007/s00799-009-0057-1>

Data Archiving and Networked Services (DANS). File formats.

<https://dans.knaw.nl/en/about/services/easy/information-about-depositing-data/before-depositing/file-formats>

Digital Preservation Strategy for the State and University Library, Denmark. Version 4, February 2016. <https://en.statsbiblioteket.dk/national-library-division/digital-preservation-strategy>

Digital Repository of Ireland Preservation Policy. March 23, 2018.

<https://repository.dri.ie/catalog/zw13bm274>

ICPSR Digital Preservation Policy Framework. Version 4, August 13, 2018.

<https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/preservation/policies/dp-p-framework.html>

Illinois Data Bank. *Illinois Data Bank Policy Framework and Definitions*.

<http://hdl.handle.net/2142/91039>

National policy for research data management in Norway (12/2017).

<https://www.regjeringen.no/contentassets/3a0ceaaa1c9b4611a1b86fc5616abde7/no/pdf/f-4442-b-nasjonal-strategi.pdf> (p. 20, in Norwegian)

Odum Institute Data Archive Digital Preservation Policy. Issued May 1, 2017.

https://odum.unc.edu/files/2017/05/Policy_DigitalPreservation_20170501.pdf

Preservation Plan Data Archiving and Networked Services (DANS). Version 1.0 – May 2018.

<https://dans.knaw.nl/en/about/organisation-and-policy/policy-and-strategy/preservation-plan-data-archiving-and-networked-services-dans-1>

Reference Model for an Open Archival Information System (OAIS), Recommended Practice, CCSDS 650.0-M-2 (Magenta Book) Issue 2, June 2012.

<http://public.ccsds.org/publications/archive/650xom2.pdf>

RLG/OCLC Working Group on Digital Archive Attributes: Trusted Digital Repositories: Attributes and Responsibilities, 2002.

<https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>

Qualitative Data Repository (QDR) Preservation Policy.

<https://qdr.syr.edu/policies/digitalpreservation>

Strategic plan for UiT The Arctic University of Norway 2014-2022. https://en.uit.no/om/art?p_document_id=377752&dim=179033

Sustainability of Digital Formats: Planning for Library of Congress Collections.

<https://www.loc.gov/preservation/digital/formats/fdd/descriptions.shtml>

The National Archives. The technical registry PRONOM.

<http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

The NDSA Levels of Digital Preservation: An Explanation and Uses.

https://ndsa.org/documents/NDSA_Levels_Archiving_2013.pdf

The Ohio State University Libraries' Digital Preservation Policy Framework.

https://library.osu.edu/documents/SDIWG/Digital_Preservation_Policy_Framework.pdf

UK Data Archive Preservation Policy. Version 09.00, 15 June 2016. <https://www.data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf>

University of Glasgow Digital Preservation Policy. Version 3.7.

https://www.gla.ac.uk/media/media_598622_en.pdf

University of Victoria Libraries Digital Preservation framework. Version 29 March 2017.

<https://www.uvic.ca/library/featured/digitalpreservation/dp-framework-FINAL.pdf>

Wilson, Andrew. Significant Properties of Digital Objects, Andrew Wilson, National Archives of Australia, 2008. <https://www.dpconline.org/docs/miscellaneous/events/142-presentation-wilson/file>

York University Digital Preservation Implementation Plan.

<https://digital.library.yorku.ca/documentation/digital-preservation-implementation-plan>

York University Digital Preservation Strategic Plan.

<https://digital.library.yorku.ca/documentation/digital-preservation-strategic-plan>

York University Environmental Monitoring of Preservation Formats.

<https://digital.library.yorku.ca/documentation/environmental-monitoring-preservation-formats>

Contact researchdata@hjelp.uit.no with questions or to request an addition or revision to this policy.

Policy Document History and Version Control Table

Version	Action	Approved By	Action Date
3.1	Policy updated.	Board of DataverseNO	2020-03-30
<u>3.0</u>	Policy revised.	Board of DataverseNO	2019-10-03
<u>2.0</u>	Policy revised.	Board of DataverseNO	2019-03-06
<u>1.0</u>	Policy issued.	Board of DataverseNO	2018-06-21