Application nestor Seal for Trustworthy Digital Archives

By Data Archiving and Networked Services (DANS) in the Netherlands
30 September 2015

Repository: The Archive of Data Archiving and Networked Services (DANS)

Address Institute:

Data Archiving and Networked Services (DANS)
Anna van Saksenlaan 51 | 2593 HW Den Haag
P.O. Box 93067 | 2509 AB Den Haag
The Netherlands
+31 70 349 44 50 | info@dans.knaw.nl | dans.knaw.nl

Contact:

- Ingrid Dillo - Deputy director
  ingrid.dillo@dans.knaw.nl
  +31 6 12 16 69 89

- Elly Dijk – Policy advisor
  elly.dijk@dans.knaw.nl
  +31 6 23 29 73 81

Table of Contents

Introduction

The Archive of Data Archiving and Networked Services (DANS) is the subject of this self-assessment. DANS is the national Dutch organization for permanent access to research data, with a focus on the humanities and the social sciences. The mission of DANS is to promote sustainable access to digital research data. For this purpose, DANS encourages researchers to archive and reuse data in a sustainable manner, e.g. through the online archiving system EASY. EASY has been developed for self-archiving.

In September 2015 there are over 30,000 datasets in the Archive of DANS. The last 5 years the average growth was nearly 4,000 datasets per year.

DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and Netherlands Organisation for Scientific Research (NWO).

The name "EASY", which occurs in some of the documents sent with the self-assessment, is identical to "the Archive" used in the other documents.

C1 Selection of information objects and their representations

Criteria have been defined for the selection of information objects and their representations in the digital repository. The framework is provided by legal obligations, the institution's or company's basic function, its own targets.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Selection Policy
For DANS it is crucial that research data remain accessible and usable in the long term. This concerns data generated or captured in academic research in any domain, with a focus on the humanities and the social sciences: the Archive's designated communities. Rather than actively setting out to acquire specific data, the Archive's policy is to leave it to the researcher and the communities to decide which data are eligible for long-term archiving and access. This approach is supported by outreach efforts ranging from consultancy, data acquisition staff liaising with specific academic disciplines and a biannual data prize to providing guidelines for the selection of information objects
See Document D2.6 *DANS Selection Policy*

Legal Policy
DANS is an institute of KNAW and NWO and is governed by the "Cooperation Agreement (*Samenwerkingsovereenkomst*) DANS" between NWO and KNAW from 2005, which was updated in January 2015 (see also Document D1.4). DANS is, administratively speaking, part of the KNAW, which is a legal entity. DANS is therefore not a legal entity on its own.
Information about the legal policy of DANS can be found in Document D1.10. *DANS Legal Policy*

Mission and Strategy
In Sharing data together - DANS strategy policy 2015-2020 the mission, target groups and strategy of DANS are described. See Document D1.2. for a link to the Summary of DANS strategy policy. The full text of the DANS strategy policy is also available in Dutch on the [DANS website](#).

Guidelines for appraising and selecting research data

In collaboration with 3TU.Datacentrum and SURF, DANS provides the publication "Selection of research data - Guidelines for appraising and selecting research data" as well a checklist on selecting research data. See Document D3.3. for a link to this publication.

C2   Responsibility for preservation

The digital repository assumes responsibility for the long-term preservation of the information objects on the basis of legal requirements or its own objectives. Long-term preservation means ensuring the long-term usability of the information contained in the representations.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Mission and Strategy
In Sharing data together - DANS strategy policy 2015-2020 the mission, target groups and strategy of DANS are described. See Document D1.2. for a link to the Summary of DANS strategy policy. The full text of the DANS strategy policy is also available in Dutch on the [DANS website](#).

Preservation Strategy
The mission of DANS is to promote and provide long-term access to digital research data. In order to validate results and enable re-use, the data should be discoverable, usable, reliable, citable and ideally openly available to the scientific community. It is not possible to achieve this without preservation, and therefore digital preservation is the core business of DANS.

The Preservation Strategy is an elaboration of the Preservation Policy, which among other things describes the retention period (See Document D2.1 *DANS Preservation Policy*). The Preservation Strategy outlines the framework in which the Archive operates. It starts with characterizing the Archive's content with an eye to preserving the data. Next, it describes preservation principles and migration as the preservation strategy of choice. A selection of community watch activities and bespoke services, as well as an overview of recurring monitoring processes conclude the Preservation Strategy. (See document D2.2 *DANS Preservation Strategy*).

Licence Agreement
The *Licence Agreement* (Document D2.3.) assigns the authority to the Archive to perform her task.

C3   Designated communities

The digital repository has defined its designated community/communities. This includes knowledge of the specific requirements of the designated communities which influence the selection of the services to be provided. If the designated communities or their requirements change over time, the digital repository should adapt accordingly.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Community watch and dedicated services
The designated community of the Archive consists of scholars in the humanities and social sciences. The Archive monitors the community through substantial contacts, for instance during data acquisition and ingest, in applied research projects, membership of European Research Infrastructures, pilot studies with data producers, via training & consultancy and by offering discipline-specific services. The pertaining section in the DANS Preservation Strategy describes a selection of activities and services for specific communities, whereas section "Recurring processes" provides an overview of generic processes for monitoring and improving the quality of the Archive.

See for more information Document D2.2. *DANS Preservation Strategy*

C4   Access

The digital repository ensures that authorised users in the designated communities can access the representations. This includes appropriate search possibilities. The digital repository openly declares its conditions of use and any costs which may arise, listing these in a transparent manner.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Access to the DANS Archive
The DANS Archive can be accessed online via https://easy.dans.knaw.nl.
The datasets can be found in EASY by searching the metadata or by browsing the various research disciplines. No authorisation is needed for searching or browsing metadata. The result will be a list of datasets, described by metadata, which is always visible to anyone, independently from Open or Restricted Access regimes. When datasets are downloaded the users are bound to the DANS General Conditions of Use. For a selection of datasets, files are available via order forms that are published with the dataset in the Archive. DANS does not charge for access and use of the data sets, but for the data storage, organising documentation, and consultancy.

Below is more information on the following aspects related to access:
a) Metadata search
b) Conditions of Use
c) Data request forms
d) Cost Model

a) DANS Metadata search
The DANS Archive has a search function directly available in the header of the website. See https://easy.dans.knaw.nl/ui/home. Alternatively, an 'Advanced search' can be done to target specific Dublin Core metadata fields: https://easy.dans.knaw.nl/ui/advancedsearch.

A Browse feature is additionally available: https://easy.dans.knaw.nl/ui/browse.

See for more information Document D4.8. *DANS Metadata search*

b) Conditions of Use
The rules for using data from the Archive, in conjunction with the licence conditions, are stated in the DANS Conditions of use (pdf), see Document D2.4. *DANS Conditions of Use*. These conditions also take into account the legal framework of DANS.

c) Data request forms
Most datasets in the DANS archive will contain published files, which can be downloaded/accessed by users if the access conditions are met. For a selection of datasets, files are alternatively available via order forms published with the dataset in the archive. This concerns:
– datasets with privacy-sensitive data for which specific Statements need to be signed;
– datasets with very large volumes of data, where system download restrictions would make direct downloading from the archive unpractical

See for the different forms in use by DANS Document D3.4. *Data Request Forms*

d) Cost Model

DANS does not charge for the access to the metadata or the use of the data sets, but will charge for three services.

We foresee an exponential growth of both the volume and the complexity of the data that we need to preserve in our Archive in the coming years. Our structural budget will not keep pace with this growth. A business model in which all data storage will remain free forever is not viable. Therefore DANS has developed a pricing structure.

See Document D1.8 for the *DANS Cost Model*.

C5   Interpretability

The digital repository has defined measures to ensure the long-term interpretability of at least one of the representations, thereby meeting a basic precondition for appropriate use now and in the future. This includes the interpretability of both content data and metadata. In ensuring this, the digital repository should take the needs of its designated community/communities into account. Changes to the technical environment or the designated community or communities can influence the interpretability of the objects. Using appropriate procedures, the digital repository should therefore check at regular intervals whether the objects can still be interpreted by the designated community or communities.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Questions by NESTOR:
a)   What is the status of the Preservation Strategy (Document D2.2) ... is it approved, implemented?
Answer by DANS: This document describes the processes as being performed by the archive, as such the strategy is implemented.

b)   So you have a preservation plan in detail, which refers to your situation, your assets, your stakeholder (especially for migration as a measure)?
Answer by DANS: The archive does not yet have a concrete plan or planning for the activities as mentioned in §2.4 of the Preservation Strategy (D2.2)
Last week we noticed that unfortunately the PDF version of the Preservation Strategy (D2.2) was incomplete. The correct version is attached, with apologies.

c)   Do you have evidence of undertaken preservation measures, especially format migration or an explanation, why this wasn´t necessary by now regarding the preservation plan?
Answer by DANS: Format migration is done at the ingest phase of the archival process (see D5.1 and D5.2). The designated preferred formats used are monitored (see §2.4 of the Preservation Strategy, #4). Since the archive does not exist that long, we have never had to migrate formats after the ingest phase; all current archived data are in preferred formats.

Statement

Long-term interpretability

To ensure long-term interpretability DANS undertakes various actions:

a) Prevention of file format obsolescence

Precautions can be taken to prevent file format obsolescence. One such measure is to select file formats which have a high chance of remaining usable in the far future. The Archive maintains a list of these so-called preferred formats.

During the data ingest stage the Archive converts data that is submitted in a non-preferred file format to a preferred file format. In this case the originally submitted source format is archived but will not be made accessible. The DANS Licence Agreement (see Document D2.4) gives the Archive the right to modify a format for the sake of sustainability, among other things.

See Document D3.2. *DANS Preferred Formats* for a link to the list of preferred formats.

b) Digital preservation

Digital preservation involves the retention of both the information object and its meaning, as stated in the significant properties. It is therefore necessary that preservation techniques can understand and recreate these to ensure the authenticity and accessibility of the data. Out of the main preservation strategies – information migration, technology emulation and encapsulation – the Archive has selected the information migration strategy, because the information content is considered more important ("significant") than its look and feel. Migration focuses on the preservation of the intellectual content. To preserve the data the Archive migrates the original file format to an open format that is independent of the particular hardware and software that were applied to create them. File formats that are non-proprietary, or that are proprietary but in widespread use, will tend to retain the best chance of being readable in the future. Non-standard proprietary formats, used only by a specific software program or specific software version, are likely to present problems for future use.

As written before the designated community of the Archive consists of scholars in the humanities and social sciences. The Archive monitors the community through substantial contacts, for instance during data acquisition and ingest, in applied research projects, membership of European Research Infrastructures, pilot studies with data producers, via training & consultancy and by offering discipline-specific services.

See Document D2.2. *DANS Preservation strategy*.

c) Data processing

The short version of the Data Processing Protocol outlines the processing of data that have been submitted to the DANS archive and takes two perspectives:

1. data depositor perspective: this is described in a publicly available document.
2. archivist perspective: the archivist's workflow is illustrated by a screenshot from the archiving application. The steps constituting the workflow are translated into English.

See Document D5.1. *DANS Data Processing Protocol (short version)* for the processing of data by the archivist.

In addition to the short version of the Data Processing Protocol the Archive has two internal protocols for the archivists. Both documents are in Dutch and available on request.

See Document D5.2 *DANS Data Processing Protocol (long version)* for a description of the protocols.

C6　Legal and contractual basis

The digital repository's ingest, archiving and access procedures are based on legal or contractual regulations concluded with the producers. The nature and scope of the delivery is regulated, as are the digital repository's archival obligations, the conditions of use and, where applicable, the costs.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
DANS Legal Policy
Information about the legal policy of DANS and the relevant legislations on which this policy is based can be found in Document D1.10. *DANS Legal Policy*

Licence Agreement
The *Licence Agreement* (Document D2.3.) assigns the authority to the Archive to perform her task.

C7   Legal conformity

The digital repository monitors and documents conformity with relevant regulations concerning the ingest, archiving and use of digital objects. These include: data protection, protection of the rights of affected parties, confidentiality regulations, copyright and usage rights, internal and external compliance.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
DANS Legal Policy
Information about the legal policy of DANS can be found in Document D1.10. *DANS Legal Policy*

Security Officer and Legal Advisor
Document D1.6. *DANS Job descriptions* gives a description of the tasks of the Security Officer and Legal Advisor.

DANS Declaration of Confidence for Employees
All DANS staff – including guest researchers, trainees et cetera – are obliged to sign the "Declaration of Confidentiality for Employees" in which the employee (et cetera) states that he/she will observe and maintain the utmost secrecy with regard to all confidential information that is supplied or will be supplied to him/her by DANS or by persons designated by DANS.
See Document D5.4. *DANS Declaration of Confidence for Employees*. This document is available on request.

C8   Funding

Valid budget planning exists, as does a long-term funding plan for the digital repository.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Budget Plan and long-term funding
DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and Netherlands Organisation for Scientific Research (NWO). In order to fulfill her task to provide sustained access to digital research data in the DANS Archive, DANS receives structural lump sum financing from KNAW and NWO.

See document D1.7 *DANS Budget plan* for the Multi-annual Budget DANS 2014 – 2020.

C9   Personnel

Sufficient numbers of appropriately qualified staff are available. Updated job descriptions exist which set out the required qualifications of the digital repository personnel and contain an organisational chart and/or a staff development plan based on the tasks and objectives of the digital repository.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Personnel of the Archive
The tasks and objectives of the Archive can be viewed the *Licence Agreement* (Document D2.3) and in the DANS *Conditions of Use* (Document D2.4).

Document D1.6. *DANS Job descriptions* gives an organisational chart of DANS and an overview of the Archive-related jobs. Job descriptions are available on request.

DANS uses Personal Development Plans (POP: persoonlijk ontwikkelingsplan) for the development of the staff.

C10 Organisation and processes

The organisational structure should be appropriate for the objectives, tasks and processes of the digital repository. The structural and procedural organisation should be defined. The responsibilities should be established. The digital repository is incorporated at the appropriate point in the schedule of responsibilities.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Statement
Job and roles descriptions
We distinguish six types of work within the Archive: data acquisition, ingest, process & preservation, archive management & quality control, IT Archive, and access. The staff related to the Archive is divided over five roles: data acquisitioner, IT developer, support & operations, data manager, and manager. These roles are based on the Dutch University Job Classification.

The three roles of Security Officer, Legal Officer and Preservation Officer refine this classification. They are assigned to employees via a letter from the DANS director.

See Document D1.6. *DANS Job descriptions* for an overview of the Archive-related jobs and roles.

See Document D1.8 for the *DANS Cost Model* for the relation with the funding.

C11 Preservation measures

The digital repository should conduct strategic planning as a means of preserving the digital objects entrusted to it. This should include imminent or expected tasks and specify the deadlines by which they are to be completed. Long-term planning should be based on the monitoring of legal and social changes, the demands and expectations of the designated communities and all technical changes relevant for the sustained preservation and appropriate use of the information objects in the form of their representations. Possible effects on task fulfilment are evaluated. Suitable structures and procedures exist for this.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

See C5 for questions by NESTOR and answers by DANS on this subject.

Statement
Preservation Policy and Preservation Strategy
The mission of DANS is to promote and provide long-term access to digital research data. In order to validate results and enable re-use, the data should be discoverable, usable, reliable, citable and ideally openly available to the scientific community. It is not possible to achieve this without preservation, and therefore digital preservation is the core business of DANS.
DANS has two documents about preservation: the Preservation Policy and the Preservation Strategy. The *DANS Preservation Policy* (See Document D2.1 for the link) details the commitment to support the long-term management of data and also outlines the roles and responsibilities of all those involved in the collection and management of data.
The *DANS Preservation Strategy* (Document D2.2) is an elaboration of the policy and outlines the framework in which the Archive operates.

Related documents are:
Document D3.2. *DANS Preferred formats* (for Preservation Policy, Paragraph 6)
Document D1.6. *DANS Job Descriptions* (for Preservation Policy, Paragraph 5)

C12 Crisis / successorship management

The digital repository is in possession of a plan which ensures continuation of the preservation tasks even beyond the repository's own existence. The digital repository should have made contingency plans. In such a case the preservation work must be continued in a different organisational framework, thereby ensuring that the set tasks can be carried out in full. Where this is not possible, any deficiencies should be documented. The digital repository should take precautions to ensure that the transition process can be defined, planned and implemented in good time.

To what extent must the criterion be met?
Implemented, 10 points

Self-assessment by DANS
Implemented, 10 points

Question by NESTOR:
Concerning C12 we are wondering, if it is possible to migrate information packages (metadata, objects) from EASY to a successor system without loosing any information.

Answer by DANS:
It is indeed possible to migrate information packages to a successor system without loosing any information.

Statement
Contingency plans
DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and Netherlands Organisation for Scientific Research (NWO). In the event of termination of the agreement NWO and KNAW will make arrangements regarding the discontinuity or modified continuity of DANS. In that case KNAW and NWO will accept in particular the responsibility to store the data files archived at DANS elsewhere in a most responsible manner and under equivalent technical conditions.
See Document D1.4. *DANS Collaboration Agreement.*

The external service provider of the Archive, I&A, is also part of the KNAW. Regarding confidentiality the Service Level Agreement (SLA) refers to the Code of Conduct for ICT and Communication Facilities of KNAW. Both documents are available on request.
See Document D4.1. *Service Level Agreement External Storage Provider.*

The plans and procedures regarding crises are laid down in Business Continuity Plan – *Crisismanagementplan DANS 2015*. The document (in Dutch) includes information on:

- the composition, roles and responsibilities of the crisismanagement team;
- the protocol when a crisis occurs;
- information on the susceptibility of the relevant locations for the operation of the Archive.

See Document D4.2. *DANS Crisis Management Plan* for more information.

C13　Significant properties

The digital repository identifies and documents which of the transferred representations' properties are significant for preservation of the information objects. In determining the scope of the properties to be preserved, a balance should be struck, bearing the repository's own targets in mind, between the technical possibilities and the costs of long-term preservation on the one hand and the needs of the designated community/communities on the other hand.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Planned, 3 points

Assessment by NESTOR
Planned in detail, 6 points

Reaction of NESTOR:
The reviewer states that 6 points are appropriate, because the archive described in detail the way how to identify the significant properties taking into consideration the affected resources and an elaborated example is provided.

Reaction of DANS:
Regarding C13 (and C18): we are pleased with this positive feedback. DANS strives to be a front-runner on the field of preserving digital research data and sets itself high ambitions. Your feedback indicates that we are on the right track.

Statement
Significant properties
DANS currently keeps an original copy of the deposited data in order to have an authentic copy. Transformations to 'preferred file formats' are performed either by the data producer, of by the data manager. Currently the significant properties are not determined or recorded explicitly, and authenticity is ensured by the professional judgement of the data producer and/or data manager.

Document 2.5. *Significant Properties and Authenticity in the Archive* describes the vision and global plan for DANS to start establishing, recording and preserving significant properties. Its challenges are to deal with heterogeneous research data with very diverse designated communities. It intends to define reference sets of significant properties per file type, per designated community. It will

focus its efforts on the most common file types for the most common designated communities.

C14 Integrity: Ingest interface

The digital repository has its own interface for ingesting the representations in a way which retains their integrity. The interface contains all of the functions and processes aimed at transferring the submission information packages from the producers, transforming them into archival information packages and incorporating them into the digital repository. The interface allows the producer and the digital repository administration to check and maintain the integrity of the representations.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Planned in detail, 6 points

Statement
Ingest as part of the Archival Process
Integrity is discussed in *D4.5 DANS Archival Integrity*, supported by *D4.3 DANS Architecture Guide* and *D4.4 Archival Process Model*. The process model is used as a basis to describe integrity for the different activities. The architecture guide describes the technical implementation of the archive (and integrity).

Fixity during deposit is implemented during transfer phase (via HTTPS and SWORD). After deposit, fixity is calculated and recorded for each file. Potential errors can be reported to the depositor is fixity information was provided with the deposit. Otherwise, the archive returns the calculated fixity information and asks the depositor to verify these.

Planned improvements:
- Virus-scans are performed periodically. Scans are currently being implemented directly after ingest.
- Policy and procedures are not explicit. These are being reviewed and will be formulated and published in 2015-Q4.

C15 Integrity: Functions of the archival storage

The archival storage provides functions necessary for checking and maintaining the integrity of the representations by the administration of the digital repository. The functions include recording of the archival information packages onto storage media, long-term storage, restoration of the archival information packages and all changes to the packages.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Question by NESTOR:
The references in the Document D4.1SLAExternalStorageProvider are not working: the referenced document is not in the folder, we can´t access the intranet of KNAW. Please send us the SLA (or a short description of the SLA).

Answer by DANS:
Regarding the document D4.1SLAExternalStorageProvider: both referenced documents are attached for the reviewers. Please note their confidential nature.

Statement
Integrity of the representations
Integrity is discussed in *D4.5 DANS Archival Integrity*, supported by *D4.3 DANS Architecture Guide* and *D4.4 Archival Process Model*. The process model is used as a basis to describe integrity for the different activities. The architecture guide describes the technical implementation of the archive (and integrity).

Fixity for the archival storage is implemented on storage- and on object level. Data is stored and backed-up automatically (see *Document D4.1 SLA External Service Provider*), and the media are automatically monitored for block-level integrity. In addition, the Fedora system records fixity for each object (data and metadata files). These are monitored periodically. Reported errors will be resolved manually.

Planned improvements:
- Intended changes to fixity are documented manually. These will need to be done more systematically.
- Procedures for changing and documenting fixity will be reviewed, formulated and published in 2015-Q4.

- Procedures for responding to fixity errors of stored objects will be reviewed, formulated and published in 2015-Q4.

The external service provider of the Archive, I&A, is also part of the KNAW. Regarding confidentiality the Service Level Agreement (SLA) refers to the Code of Conduct for ICT and Communication Facilities of KNAW. Both documents are available on request.
See Document D4.1. *Service Level Agreement External Storage Provider.*

C16 Integrity: user interface

The digital repository has an interface, which allows users and the digital repository administration to check and maintain the integrity of the representations. This includes the transformation from archival information packages into dissemination information packages.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Planned in detail, 6 points

Statement
Integrity: user interface
Integrity is discussed in *D4.5 DANS Archival Integrity*, supported by Document *D4.3 DANS Architecture Guide* and Document D*D4.4 Archival Process Model*. The process model is used as a basis to describe integrity for the different activities. The architecture guide describes the technical implementation of the archive (and integrity).

Fixity information is currently visible after uploading files via the user interface, in the depositor-license and in the DIP. Visibility of for data managers and data consumers via the user interface is planned for 2016-Q1.

C17 Authenticity: Ingest

The digital repository has procedures, which permit the authenticity of the representations to be assessed upon being ingested and the authenticity of the submission information packages to be assessed and protected.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Ingest authenticity
During the acquisition of data, the archive recommends or enforces the data producer to submit its data using preferred- or accepted formats (See document D3.2. *DANS preferred formats* for a link to the list of preferred formats). This way the data producer is responsible for potential transformation, and as such for ensuring the authenticity of the data. In addition, the data producer should provide available documentation about the creation of the data, and about how the data can be used.

After user authentication, the data can be deposited via the web interface, via a SWORD interface or via bulk import (via e.g. harddrive, FileSender or FTP upload). There are no transformations to the data itself involved, other than packing and unpacking. The integrity of the files is ensured via checksums. Once data, metadata and appropriate rights have been transferred, the deposit is complete.

See Document D2.5. *Authenticity in the Archive* for more information.

C18 Authenticity: Preservation measures

The digital repository deploys methods, which ensure the authenticity of the objects during implementation of the long-term preservation measures and document the degree of authenticity.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Planned, 3 points

Assessment by NESTOR
Planned in detail, 6 points

Reaction of NESTOR:
Reviewer: C18 and C24: for both 6 points are appropriate (the archive scored C18: 3; C24: 10) because in both cases for the measures which take place during ingest the criteria are fully implemented, for migration they are planned.

Reaction of DANS:
Regarding (C13 and) C18: we are pleased with this positive feedback. DANS strives to be a front-runner on the field of preserving digital research data and sets itself high ambitions. Your feedback indicates that we are on the right track.

Statement
Preservation authenticity
Assessment and curation of the data take place at two events: after a dataset is deposited and when an existing datasets may no longer comply with the latest (updated) requirements of the archive.

See Document D2.5. *Authenticity in the Archive* for more information.

C19　Authenticity: Use

The digital repository allows the users and the administration of the digital repository to check and maintain the authenticity of the representations. This includes the transformation from archival information packages into dissemination information packages.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
User authenticity
The archive allows the curated data to be ordered. The data consumer will receive (after user authorization) the curated files from the archive, without transformations (other than packaging multiple files). The original files and the provenance document and/or details about the significant properties are currently not provided via the common user interface, but can be provided upon request.

See Document D2.5. *Authenticity in the Archive*

See the example of the workflow in Document D5.1. *DANS Data Processing Protocol (short version).* In the field 'remarks' the data archivist documents all the transfers of the data set.

C20 Technical authority

The digital repository obtains technical authority over the representations being ingested, allowing it to transform them into archival information packages and, if necessary, to carry out long-term preservation measures. Following the transfer, all necessary measures can be carried out without any technical restrictions.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Technical authority over the ingested representations
According to the *DANS Data Processing Protocol* (see Document D5.1 *DANS Data processing Protocol*) the archivist checks at ingest whether the submitted files are accessible and readable. Any encrypted file would be returned to the data producer.

The *Licence Agreement* (Document D2.3.) is considered as the Submission Agreement between the Archive and an individual Producer. Thus, it does not cover agreements between Archive and institutional Producers, which are laid down in separate contracts. Examples of these contracts are available upon request.

See Document D2.7. *DANS submission Agreement* for a table that shows which elements are dealt with in the different documents between the Archive and the Producers.

See Document D3.1. *DANS Depositing Instruction* website for instructions for several disciplines.

C21 Submission information packages

The digital repository has issued specifications regarding its submission information packages. The digital repository agrees with the producers which submission information packages will be ingested (content data and metadata). The submission information packages are checked on the basis of the specifications.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Specifications submission information packages
Submission information packages (SIPs) are specified to the producers at the moment of ingest.
See Document D4.9. *DANS SIP Instructions* for the three possible methods for ingesting SIPs in the Archive.

See Document D4.3. *DANS Architecture Guide* for a link to the Archival Information Package (AIP) design with information about all Information Packages (page 3).

See Document D4.4. *DANS Archival process Model* page 4. for the description of the activities within the archival processes: Deposit Data about SIPs.

C22 Transformation of the submission information packages into archival information packages

The digital repository converts submission information packages into archival information packages

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Archival information packages
The curate data activity is performed by the archive and involves the necessary actions to ensure that the data conforms to the collection policy, as identified in the assessment report. Such actions may involve format transformations and metadata enrichment, but may also include restoring data from backup (when integrity was damaged). The result is an Archival Information Package (AIP) that will be archived. Curate data is part of the processes "Archive Data" and "Preserve Data". The former will involve curation of data after their deposit, the latter will involve curation after changes in the designated community and/or the collection policy.
See Document D4.4. *DANS Archival process Model* for the description of the activities within the archival processes: Curate Data about AIPs.

See Document D4.3. *DANS Architecture Guide* for a link to the Archival Information Package (AIP) design with information about the Archival Information Packages.

The *Licence Agreement* (Document D2.3.) assigns the authority to the Archive to perform this task.

C 23 Archival information packages

The digital repository has issued specifications for its archival information packages. The digital repository defines which archival information packages (content data and metadata) are to be stored and in which form. The archival information packages are checked on the basis of the specifications.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Specifications archival information packages
See Document D4.3. *DANS Architecture Guide* for a link to the Archival Information Package (AIP) design with information about the Archival Information Packages.

C 24 Interpretability of the archival information packages

Technical preservation measures are undertaken to ensure the interpretability of the archival information packages.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

See C5 for questions by NESTOR and answers by DANS on this subject.

Assessment by NESTOR
Planned in detail, 6 points

Reaction of NESTOR:
Reviewer: C18 and C24: for both 6 points are appropriate (the archive scored C18 : 3; C24: 10) because in both cases for the measures which take place during ingest the criteria are fully implemented, for migration they are planned.

Reaction of DANS:
DANS accepts the lower rating, since here DANS is indeed somewhere between the planning and the implementation phase.

Statement
Interpretability of the archival information packages
This whole process starts with the precautions that can be taken to prevent file format obsolescence. One such measure is to select file formats which have a high chance of remaining usable in the far future. The Archive maintains a list of these so-called preferred formats.
See Document D3.2. *DANS Preferred Formats* for the list of preferred formats*.*

Out of the main preservation strategies – information migration, technology emulation and encapsulation – the Archive has selected the information migration strategy, because the information content is considered more important ("significant") than its look and feel.
The Archive's preservation processes are organised according to the OAIS reference model.
See Document D2.1. *DANS Preservation Policy*

C 25 Transformation of archival information packages into dissemination information packages

The digital repository transforms archival information packages into dissemination information packages.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Archival information packages into dissemination information packages

See Document D4.3. *DANS Architecture Guide* for a link to the Archival Information Package (AIP) design with information about the dissemination information packages.

.

C26 Dissemination information packages

The digital repository specifies the dissemination information packages on the basis of the requirements of the designated communities.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Planned in detail, 6 points.

Statement
Specification the dissemination information packages
In response to requests from the designated community DANS data managers can convert files to a desired format. Furthermore, for instance for statistical data several file formats are provided in the Archive.

See Document D4.3. *DANS Architecture Guide* for a link to the Archival Information Package (AIP) design with information about the dissemination information packages (DIP).

See Document D3.2. *DANS Preferred Formats* for a link to the list of preferred formats.

C27 Identification

A digital repository should use internal identifiers to manage the information objects and their representations and, where applicable, their parts and relationships (part/totality, different variants, versions etc.), especially to ensure unique assignment of the content data to the metadata.
The use of externally visible, standardised persistent identifiers ensures reliable tracing of the information objects and their representations, and consequently also access.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Internal and external identifiers
Datasets in the Archive can be identified by 3 kind of identifiers:
- internal Fedora identifiers for objects like datasets, files, and folders
- external URN (persistent) identifiers for long-term identification of datasets within the national repository infrastructure.
- external DOI (persistent) identifiers for citation of datasets.

See Document D4.6. *DANS Metadata* for more information on identifiers in the Archive. When submitting a dataset to the Archive the producer receives the newly minted DOI via e-mail. Also, all identifiers are visible to all users in a dataset's metadata

C28 Descriptive metadata

The scope, structure and content of the descriptive metadata are defined. They depend on the goals of the digital repository, its designated communities and the object types.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
The scope, structure and content of the descriptive metadata
The Archive contains descriptive metadata on two levels: the Dataset level and the File Item level.

See Document D4.6. *DANS Metadata* for more information on descriptive metadata in the Archive. To ensure that producers observe the description guidelines, the archivists check the submitted metadata (along with the data and other documentation), see Document D5.2 *DANS Data Processing Protocol*. This often results in contacting the producer to improve the metadata.

C29 Structural metadata

The structure of the representations must be adequately described so that the information objects can be reconstructed and used.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Structural metadata
The structure of a dataset is described in the RELS-EXT data stream of File Items and Folder Items. RELS-EXT is an RDF/XML document. The Archive defines two types of relationships between a subject and an object.

See Document D4.6. *DANS Metadata* for more information on structural metadata in the Archive.

C30 Technical metadata

The technical metadata are defined to ensure interpretability, integrity and authenticity and to manage the preservation measures.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Assessment by NESTOR
Planned in detail, 6 points.

Reaction of NESTOR:
The reviewer suggests 6 points. File-format (MIME-Type), checksums and file size are few technical metadata. We suggest considering a more elaborated technical metadata schema to be prepared better for risk analysis and migration planning.

Reaction of DANS:
DANS knows that her technical metadata are not yet very rich.

Statement
Technical metadata
Datasets in the Archive contain only one technical metadata field: Format, which has the semantics of the Dublin Core (DCES) "Format" element. It is a repeatable field that describes the formats present in the Dataset, preferably in the MIME vocabulary (MIME).

See Document D4.6. *DANS Metadata* for more information on technical metadata in the Archive.

C31   Logging the preservation measures

The digital repository logs the preservation measures and any changes to the representations.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Logging preservation measures
Workflow steps are intended to document the preservation activities undertaken by the archivists. Therefore, the name of the archivist and the time of completion are logged as soon as he checks the items as "done" in the Web-Userinterface.

See Document D4.6. *DANS Metadata* for more information on logging the preservation measures in the Archive. Changes are recorded in the ADM Metadata.

See Document D5.2 *DANS Data Processing Protocol*. The checkboxes allow (semi-automatic) recording of default preservation measures upon ingest. The remarks entry is used by archivists to (manually) record all other changes to the dataset.

C32 Administrative metadata

The digital repository has defined its administrative metadata in order to render the administration and use of the information objects and their representations transparent.
Use of the representations may be restricted for legal or contractual reasons.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Administrative metadata
The Archive contains administrative metadata (AMD) on the dataset level and on the File Item level.

See Document D4.6. *DANS Metadata* for more information on administrative metadata in the Archive.

C33 IT infrastructure

The IT infrastructure should realise the specifications for handling the information objects and representations at the technology and security levels.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Question by NESTOR:
A technical documentation on the EASY architecture based on FEDORA would help to review C33.

Answer by DANS:
We attached an export of the EASY architecture description, which includes a more detailed description of the Fedora implementation in EASY.

Statement
IT infrastructure
IT infrastructure is a strategic component of DANS to achieve its mission and provide long-term access to digital research data.

Document D4.7 *IT Infrastructure* provides an inventory of the IT infrastructure and explains how it supports the needs of the archive and the values of DANS, and how it is maintained and developed to meet todays / tomorrows requirements.

C34 Security

The organisation and the infrastructure protect the digital repository and its archived information objects and representations.

To what extent must the criterion be met?
An average of 7 points must be achieved in the assessment of the applicable criteria C13 - C34.

Self-assessment by DANS
Implemented, 10 points

Statement
Security policy
The security policy at DANS is in place to ensure that valuable digital data remains available (Availability), and does not become corrupted (Integrity), or fall into the wrong hands (Confidentiality). If data that is managed by DANS is lost or damaged it might damage confidence in DANS as data custodian. A serious breach of information security could even jeopardize the continued existence of DANS.

See Document D1.9. *Security Policy* for more information on the security of the digital repository and its archived information objects and representations.

The plans and procedures regarding crises are laid down in Business Continuity Plan – Crisismanagementplan DANS 2015. The document (in Dutch) includes information on:
- the composition, roles and responsibilities of the crisis management team;
- the protocol when a crisis occurs;
- information on the susceptibility of the relevant locations for the operation of the Archive.

See Document D4.2. *DANS Crisis Management Plan* for more information and References.