



## **Implementation of the Data Seal of Approval**

The Data Seal of Approval board hereby confirms that the Trusted Digital repository DANS: Electronic Archiving SYstem (EASY) complies with the guidelines version 2014-2017 set by the Data Seal of Approval Board.

The afore-mentioned repository has therefore acquired the Data Seal of Approval of 2013 on November 21, 2013.

The Trusted Digital repository is allowed to place an image of the Data Seal of Approval logo corresponding to the guidelines version date on their website. This image must link to this file which is hosted on the Data Seal of Approval website.

Yours sincerely,

The Data Seal of Approval Board

## Assessment Information

Guidelines Version: 2014-2017 | July 19, 2013  
Guidelines Information Booklet: [DSA-booklet\\_2014-2017.pdf](#)  
All Guidelines Documentation: [Documentation](#)

Repository: DANS: Electronic Archiving SYstem (EASY)  
Seal Acquiry Date: Nov. 21, 2013

For the latest version of the awarded DSA for this repository please visit our website: <http://assessment.datasealofapproval.org/seals/>

Previously Acquired Seals: Seal date: April 12, 2011  
Guidelines version: 2010 | June 1, 2010

This repository is owned by:

**DANS**  
Anna van Saksenlaan 51  
2593 HW Den Haag  
The Netherlands

T +31 70 349 44 50  
E [info@dans.knaw.nl](mailto:info@dans.knaw.nl)  
W <http://www.dans.knaw.nl/>

# Assessment

## 0. Repository Context

### Applicant Entry

*Self-assessment statement:*

DANS is an institute of the Royal Netherlands Academy of Arts and Sciences ([KNAW](#)) and Netherlands Organisation for Scientific Research ([NWO](#)). 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. The basic unit in EASY is called “dataset” and a person who submits a dataset is called a “depositor”. A depositor need not be identical to the data producer. Unless indicated otherwise, however, in this self-assessment the DSA term “data producer” and the EASY term “depositor” refer to same role.

As part of its mission, DANS supports the Open Access principle, while being aware of the fact that not all data can be freely available and without limitations at all times. Therefore, DANS applies the principle ‘Open if possible, protected if necessary’. It is important that all research data – whether publicly available, not available (yet) or only available to a limited degree – are archived in a sustainable manner. Both Open and Restricted access regimes are supported by licences and workflows. In [OAIS](#) parlance, the Access function is essential for EASY.

DANS/EASY does not outsource any of the DSA guidelines. However, as described under Guideline 6, data storage management has been outsourced. The Service Level Agreement with the data storage management provider is available on request.

### Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

This is a good, succinct explanation of the repository purpose and approach.

# 1. The data producer deposits the data in a data repository with sufficient information for others to assess the quality of the data, and compliance with disciplinary and ethical norms.

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

## Applicant Entry

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

DANS has an Electronic Archiving SYstem, named EASY. Depositing in a trusted digital repository, like EASY, has various positive effects on the quality of the data and the discourse about data quality. Depositing in an archive that is known in the research field stimulates the researchers to act very responsibly in regard to methods and techniques that prevail in their disciplines. It also makes researchers who perform substandard open to criticism by peers.

In the instructions for depositing (<http://www.dans.knaw.nl/en/content/data-archive/depositing-data>) DANS stimulates data producers to document their source material, research methods and publications related to the data. These instructions vary by discipline: Archaeology; History, Language and Literature Studies; Life Sciences and Medicine; Social and behavioural sciences; and "Other disciplines". The instructions describe what is obligatory and what is (strongly) recommended. The use of DDI metadata (DDI = Data Documentation Initiative, <http://www.ddialliance.org/>) is obligatory for special (longitudinal) social science research programs. Providing Dublin Core metadata (<http://dublincore.org/documents/dcmi-terms/>) is obligatory for all depositions (see also Guideline 3).

A DANS archivist evaluates each deposit, the file formats, metadata and the accompanying documentation. When necessary the archivist will contact the data producer and ask for more information. In rare cases the deposit may even be put back into draft mode, so the depositor/data producer can improve his submission. The archivists have expertise in various disciplines and have a general knowledge of the research activities that take place in those fields. In this way they can judge the quality of a deposit, with a focus on accessibility and re-usability of the data. For a detailed description of the archivists' work process at data ingest see Guideline 12.

As for the ethics, the archivists also keep a close look on the possible existence of information in the research data that could identify individual persons or could otherwise be against the law (like racism, child pornography etc.). For Open Access data this information will be removed. Whether the publication of personal data is adequately (i.e. restricted) licensed is also checked according to the DANS protocol for data archivists (see Guideline 8).

A strategic target of DANS for 2015 is that a mechanism for data quality assessment has been created on the basis of data reviews. Since 2010 DANS has been collecting “data reviews” in EASY: critical reviews by the consumers (not necessarily peers) of a given dataset in EASY. The number of reviewed datasets is growing and the review functionality is to be integrated with EASY. The data reviews are part of the metadata that allow others to assess the quality and usability of the data. See for an example:

<http://datareviews.dans.knaw.nl/details.php?l=en&pid=urn:nbn:nl:ui:13-0an-1ei>

### **Planned activities:**

DANS is exploring possibilities for joining scientific peer review processes. It should then be visible in EASY who has assessed a dataset and what that assessment resulted in. Possibly, peer review may also take place anonymously at the discretion of DANS or of an authority in the discipline to which a dataset belongs.

### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

The data review example is very interesting. In general, I really like the addition of Planned Activities throughout the self-assessment. It gives confidence that DANS is continually improving and is committed to compliance with all guidelines.

On a minor note, the first sentence of the next to last paragraph has characters running together without spaces. I tested this in both Firefox and IE and the text displays the same way (run together) in both.

## 2. The data producer provides the data in formats recommended by the data repository.

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

### Applicant Entry

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

DANS considers that the file formats best suited for long-time preservation and accessibility are the file formats which are commonly used, which have open specifications, and which are independent of specific software, developers or suppliers. However, DANS is aware that in practice it is not always possible to select formats that meet with all of these ideal attributes.

DANS, therefore, offers its depositors a list of “preferred and acceptable formats”. The preferred formats are the file formats, which DANS trusts will offer the best guarantees for usability, accessibility and robustness in the long term. Basically, DANS expects these formats to be sustainable for the longer term.

The use of acceptable formats will, for a number of reasons, be allowed in the data archive as well, but long-term preservation of these formats is uncertain. DANS therefore strongly recommends data depositors to deliver their data in the preferred format corresponding to the type of data.

DANS asks depositors whose datasets contain file types different from the formats in the list to contact the data archive. Archivists check submitted datasets for their file formats and contact the depositor, if necessary.

The list of preferred formats and acceptable formats will change over time as new formats will be developed and others will fall into disuse. The list can be found at <http://www.dans.knaw.nl/en/content/data-archive/depositing-data>. The DANS Licence Agreement between data producer and DANS states DANS’s right to modify the format of the dataset if this is necessary e.g., to facilitate the digital sustainability (see article 3b in [http://www.dans.knaw.nl/sites/default/files/file/archief/Licence\\_agreement\\_DANS\\_UK.doc](http://www.dans.knaw.nl/sites/default/files/file/archief/Licence_agreement_DANS_UK.doc)).

### Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

*Comments:*

The grid of preferred and acceptable formats is very useful in that it covers a range of data types.

### **3. The data producer provides the data together with the metadata requested by the data repository.**

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

#### **Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

DANS has developed EASY to this end. The data producer deposits the data into the system and adds the relevant metadata in the web form (<https://easy.dans.knaw.nl/ui/deposit>; registration required). How to do this is described in the instructions for depositing, which can be found at the DANS website: <http://www.dans.knaw.nl/en/content/data-archive/depositing-data> (see also Guideline 1). The web form also provides brief explanations and examples.

DANS follows as much as possible the specifications of Qualified Dublin Core (<http://dublincore.org/documents/dcmi-terms/>). Some metadata fields are obligatory, while the others are optional. DANS has considered to make more fields obligatory, but has decided against it on the grounds that metadata might then become a threshold resulting in researchers not offering their data. As a general policy, the number of obligatory fields is therefore kept as low as possible.

Similarly to the instructions for depositing, a data producer can select a discipline. Because the situation regarding archaeology strongly differs from the approach to other disciplines, both metadata situations are described here.

Generically, the obligatory metadata fields are Title, Creator, Date created, Description, Access rights, Date available, Audience. The other fields are recommended (Contributor(s), Subject, Spatial coverage, Temporal coverage, Source, Identifier) or presented as additional (Format, Relation, Language, Remarks). Whenever possible, drop-down menus are provided for efficiency and consistency.

For the e-depot for Dutch archaeology (EDNA), accommodated at DANS and part of EASY, there is a number of extra fields: Archis research registration no., Alternative title, (Copy)Rightholder, Publisher. In addition a few fields, such as Spatial coverage, are more detailed for archaeology.



To accommodate the international research infrastructure of CLARIN (<http://www.clarin.eu/>), in which DANS participates, the web form for language and literature depositions includes an extra check on the availability of dedicated CLARIN metadata.

Such discipline-specific metadata (cf. the DDI metadata mentioned in the context of Guideline 1), as well as the focus in instructions and personal communication with data producers on rich metadata is aimed to improve data discoverability with its ensuing benefits for data producers, such as higher visibility.

**Planned activities:**

DANS expects that more discipline-dependent (domain-dependent) metadata fields will be defined in EASY, notably within the framework of international research infrastructure projects in which DANS is involved; see Guideline 10 for more examples.

**Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

#### **4. The data repository has an explicit mission in the area of digital archiving and promulgates it.**

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

#### **Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

The mission of DANS and its strategy policy for the period 2011-2015 can be found at the DANS website:

<http://dans.knaw.nl/content/strategie-en-beleid>. The English summary of the DANS strategy policy 2011-2015 is available at:

[http://www.dans.knaw.nl/sites/default/files/file/jaarsverslagen%20en%20strategienota/Samenvatting%20strategienota\\_UK\\_DEF.pdf](http://www.dans.knaw.nl/sites/default/files/file/jaarsverslagen%20en%20strategienota/Samenvatting%20strategienota_UK_DEF.pdf).

(Last seen 29 October 2013)

DANS is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW, [https://www.knaw.nl/en?set\\_language=en](https://www.knaw.nl/en?set_language=en)) and Netherlands Organisation for Scientific Research (NWO, <http://www.nwo.nl/en>). It has been founded and is structurally funded with the explicit mission 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. 'Digital research data' is: research information, research data (such as databases, spreadsheets, text, images, audio, video, multimedia) and digital publications (including preprints, reports). DANS supports its services with training and advice, as well as research into sustainable access to digital information. Central to sustainable data storage is that the data should be traceable, accessible and usable at all times. The Data Seal of Approval (see: <http://datasealofapproval.org/en/>), which was originally developed by DANS, is used as a criterion for this, regardless of whether the data are stored at DANS or elsewhere.

**Following from this mission DANS has formulated four strategic priorities:**

- DANS will strengthen its services by serving more users more efficiently.

DANS offers services that are tuned to the demand. See for an overview the Dutch page

(<http://www.dans.knaw.nl/content/diensten>) or a limited overview in English

(<http://www.dans.knaw.nl/en/content/services>).

- DANS will develop into a discipline-independent data organisation.

Until 2011 DANS focused on the humanities and social sciences, but some services, e.g. persistent identifiers (see Guideline 10) are discipline-transcending. Therefore DANS now intends to serve all disciplines, by first focusing on those areas where there is an explicit demand for data services.

- DANS will conduct research to support and improve its services.

DANS has established an e-Research Group that focuses on the support and innovation of the primary services of DANS. See for the e-Research Programme “Exploring the Long Term Availability of Research Data”:  
[http://www.dans.knaw.nl/sites/default/files/file/DANSeResearchProgramme\\_Draft.pdf](http://www.dans.knaw.nl/sites/default/files/file/DANSeResearchProgramme_Draft.pdf)

- DANS will be an important building block in data provision in Europe.

DANS participates in various (inter)national research and data infrastructures (see Guidelines 3 and 10).

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

This detail about strategic goals augments the mission and is a useful addition.

One thing that is not prominent in the evidence is succession planning, but it seems likely that DANS is in discussions with other possible partners in the Netherlands and Europe about this topic. The rating of "4" is still appropriate.

## **5. The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.**

### *Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

## **Applicant Entry**

### *Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

### *Self-assessment statement:*

DANS is an institute of KNAW and NWO and is governed by the “Cooperation Agreement (Samenwerkingsovereenkomst) DANS” between NWO and KNAW from 2005(see also Guideline 4). Legally it is part of the KNAW, which is a legal entity. DANS is therefore not a legal entity on its own. See <http://www.dans.knaw.nl/sites/default/files/file/algemeen%20beleidskader%20DANS%281%29.pdf> (in Dutch).

At its website DANS refers to all relevant legal information and its consequences for depositing or distributing data, see <http://www.dans.knaw.nl/en/content/data-archive/legal-information>. (Seen 29 October 2013).

Both to depositing and using data, agreements apply: the DANS Licence Agreement and the DANS General Conditions of Use. These agreements are based upon the principles of [Open Access](#) and the relevant legislation of The Netherlands and the European Union as well as the codes of conduct for scientific research of the Dutch Association of Universities (VSNU) (see Guideline 15 for more information). An outline of the applicable legislation:

- Databases Act and Copyright Act: Datasets almost always fall within the scope of the Databases Act (Databankenwet) and sometimes within that of the Copyright Act (Auteurswet). See <http://www.dans.knaw.nl/en/content/copyright-and-publications> (Seen 29 October 2013)

- Personal data protection: It may occur that personal data are incorporated in a dataset. These data may only be processed, stored and used in accordance with the Data Protection Act (WBP) of The Netherlands. See <http://www.dans.knaw.nl/en/content/privacy-sensitive-data> as well as the privacy regulations: <http://www.dans.knaw.nl/en/content/dans-privacy-regulations>. (Seen 29 October 2013)

- Audio and visual data: Audio and visual datasets may often contain personal data. Furthermore, portrait right may apply to visual data. This concerns the identifiable representation of persons. For more information see <http://www.dans.knaw.nl/en/content/audio-and-visual-data>. (Seen 29 October 2013)

All DANS staff – including guest researchers, trainees et cetera – have signed 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. This policy is available on request.

Regarding the way that DANS supports and enforces that data consumers respect such regulations, see Guideline 16.

To ensure that the relevant legal knowledge remains up-to-date one member of staff has been given the explicit task to monitor the developments in this field and report possible changes to the director of DANS.

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

## **6. The data repository applies documented processes and procedures for managing data storage.**

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

### **Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

Work on making the Preservation Policy explicit is in progress; this policy should be final and published at the DANS website early 2014.

Late 2011 DANS has carried out an initial risk analysis on the basis of the DRAMBORA methodology (<http://www.repositoryaudit.eu/>). The Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) provides a methodology for self-assessment. This analysis led to the identification and description of a number of risks, all of which were assigned a weight and a "Risk Management Activity Owner". At the second assessment early 2013 some risks had effectively been dealt with and in many other cases the risk severity could be reduced. DANS will continue annual DRAMBORA self-assessments.

Data storage management has been outsourced. Physical data storage has been outsourced in turn by the data storage management provider. DANS has a Service Level Agreement (SLA) with its data storage management provider, which includes a confidentiality statement; this SLA is available on request. The data storage management provider in turn has an SLA with the physical data storage provider.

The data is stored on a dedicated server and every 24 hours tape backups are made which are stored in two physically separate locations. The monitoring of the servers and tape backups is automated. The storage array for the original datasets and the tapes are replaced at regular intervals. Access is restricted to engineers who perform a fixed and specified number of maintenance roles. No data recovery has been necessary so far.

In 2012 DANS has conducted a technology vulnerability scan, which revealed some low risk level vulnerabilities, which have been subsequently fixed. Subsequent scans will be conducted in future.

**Planned activities:**

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

- Finish and publish the Preservation Policy early 2014.
- Improve file fixity checking by adding checksums and checking these checksums (see also Guideline 11).
- Continue the DRAMBORA risk assessment in March 2014

### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

## 7. The data repository has a plan for long-term preservation of its digital assets.

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

### Applicant Entry

*Statement of Compliance:*

3. In progress: We are in the implementation phase.

*Self-assessment statement:*

DANS maintains a list of preferred and accepted file formats for use during the ingest phase (see also Guideline 2; <http://www.dans.knaw.nl/en/content/data-archive/depositing-data>). This list is reviewed every year and updated when necessary. The files in the archive are being monitored on a regular basis. If it looks like certain formats will become obsolete, these files are migrated to a new (preferred) format. This strategy has proven successful as files which have been stored over forty years ago can still be read.

DANS has appointed a Technical Archivist who is the main responsible for implementing the long-term preservation plan.

### Planned activities:

DANS is in the process of working out the details of a long-term preservation plan based on its policy. This plan should be finalised in 2014.

### Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)



The ranking of "3" seems appropriate here. Again, having a section on planned activities gives the sense that this will move forward and likely into full compliance.

## 8. Archiving takes place according to explicit work flows across the data life cycle.

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

### Applicant Entry

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

DANS has an extensive set of instructions (see Guideline 1 and <http://www.dans.knaw.nl/en/content/data-archive/depositing-data>) (seen 29 October 2013), which describe the steps the archivists take when handling deposits. These work processes have been developed over nearly fifty years. They are described in documents and implemented in the archival workflow, which is based upon the OAIS reference model. Because the EASY system is based upon the principle of self-archiving, the first step in the process is carried out by the data producer. The data producer creates the SIP (Submission Information Package) according to the DANS instructions. The archivist then checks the metadata, the privacy clauses, and the file format. When the SIP becomes an AIP (Archival Information Package) preservation procedures ensure its readability over time. All actions related to the preservation of the data are documented in the Provenance document; see [http://www.dans.knaw.nl/sites/default/files/file/Provenance\\_document\\_120823\\_UK.pdf](http://www.dans.knaw.nl/sites/default/files/file/Provenance_document_120823_UK.pdf) (seen 29 October 2013), as well as the information regarding Guideline 12.

EASY contains various types of data; see “digital research data” listed at Guideline 4 as well as Guideline 2 for the list of preferred and acceptable formats at <http://www.dans.knaw.nl/en/content/data-archive/depositing-data> (seen 29 October 2013). Where different formats have an impact on the workflow, this is described in the various instructions. DANS may refer data producers to other Trusted Digital Repositories when these are better able to accommodate specific (discipline-dependent) types of data.

Data distribution and controlled access are done in an automated way, documented in two documents: the DANS Licence Agreement and the DANS General Conditions of Use (see also Guidelines 5 and 14, as well as <http://www.dans.knaw.nl/en/content/data-archive/legal-information>) (seen 29 October 2013).

### Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

## **9. The data repository assumes responsibility from the data producers for access and availability of the digital objects.**

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

### **Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

Because the EASY system is based upon the principle of self-archiving, both the data producer and DANS have responsibilities. These are described in the instructions for depositing and in the DANS Licence Agreement; see Guidelines 1 and 5.

EASY has two access categories: Open access and Restricted access. They are described in the context of Guideline 14.

To provide optimal access to the digital objects DANS follows a fixed procedure if EASY is interrupted – whether scheduled or not. In the latter case, when EASY should break down, users are informed by a warning page. Depending on the nature of the problem, DANS can use the audit log to trace back data producers who were depositing data at the moment of the interruption. By contacting them potential data loss can be prevented. In less severe cases of interruption EASY can be accessed in read-only mode, which precludes ingest, but allows for data access including dissemination.

In case of scheduled interruptions, for instance for maintenance, a banner on each EASY web page announces the interruption. Furthermore, during maintenance an information page is being presented.

**Planned activities:**

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

The procedures and responsibilities for dealing with interruptions will be made explicit in a crisis management plan, to be finalised in 2014.

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

## 10. The data repository enables the users to discover and use the data and refer to them in a persistent way.

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

### Applicant Entry

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

As re-using research data is key to DANS's mission, data discovery is supported in various ways. The datasets can always be found in EASY by browsing the various research disciplines or by searching the metadata, and by reading the metadata, which is always visible to anyone, irrespective of Open or Restricted Access regimes. DANS stimulates adding as many metadata and other relevant documentation as possible; see Guideline 3.

DANS participates in NARCIS, "the gateway to scholarly information in the Netherlands" (<http://www.narcis.nl/?Language=en>) to increase data discovery and to relate researchers, research data, and publications. By means of the OAI-PMH protocol (<http://www.openarchives.org/pmh/>) not just NARCIS, but anyone can harvest the metadata in EASY. This is done for instance by the research infrastructure CLARIN to present datasets that are relevant to their research community in the CLARIN portal (see also Guideline 3). Furthermore, DANS opens up EASY for searching the metadata via Google.

DANS participates in different research infrastructures like CLARIN, CESSDA (<http://www.cessda.org/>) and DARIAH (<http://www.dariah.eu/>). Besides, the metadata is reused in different portals and data services like NARCIS, CARARE (<http://www.carare.eu/>), Engage (<http://www.engagedata.eu/>), Europeana Cloud (<http://pro.europeana.eu/web/europeana-cloud>) and other national and international (topical) services.

When datasets are downloaded the data consumers (users) are bound to legal regulations, codes of conduct (see Guideline 5) and the DANS General Conditions of Use. These conditions state for instance that a user, who creates a dataset with the aid of an EASY dataset, has to deposit this new dataset in EASY.

In order to be able to refer to datasets, EASY generates a persistent identifier with the syntax "urn:nbn:nl:ui:13-xxx-yyy", where "urn:nbn:nl:ui:13-" is a fixed string that indicates the EASY system and "xxx-yyy" is a variable with a unique value for each dataset. To retrieve the particular dataset on the internet, a resolver can be used at <http://persistent-identifier.nl/>; the full link then becomes <http://persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-xxx-yyy>. The persistent identifier is visible in EASY at a dataset's tab pages "Overview" and "Description". The DANS General Conditions of Use state how datasets must be referred to in publications and how the persistent identifier must be part of this reference.

#### Data Seal of Approval Board

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

**Planned activities:**

Enabling the possibility of Open Access without the obligation of registration.

**Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

## 11. The data repository ensures the integrity of the digital objects and the metadata.

### *Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

### **Applicant Entry**

#### *Statement of Compliance:*

3. In progress: We are in the implementation phase.

#### *Self-assessment statement:*

To verify data integrity (file fixity) DANS is developing a checksum procedure for bit preservation. The related process includes technology watch and monitoring obsolescence in a way that is further automated than it currently is.

Replicability of studies is an important instrument in scientific integrity and progress. In this context it is essential that a persistent identifier always refers to the same contents. Once data have been deposited in EASY it is impossible for the data producer or for anyone except archivists to change the data. DANS distinguishes between two forms of alteration after ingest. On the one hand, when there is a change to data, this results in a new version and therefore a new dataset with its own persistent identifier (see Guideline 10). The new and the previous dataset are cross-referenced in their respective descriptive metadata. On the other hand, when there is a change to metadata, descriptive documents or supplementary files, this is considered as a minor change, which does not lead to a new dataset.

#### **Planned activities:**

- Improve file fixity checking by adding checksums and checking these; document the process and the responsibilities in 2014. The data producers will receive checksums of the files they have deposited.
  
- Finish and publish the Preservation Policy early 2014.

### **Reviewer Entry**

#### *Accept or send back to applicant for modification:*

Accept

#### *Comments:*

#### **Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

The rating of "3" seems appropriate here.



## 12. The data repository ensures the authenticity of the digital objects and the metadata.

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

### Applicant Entry

*Statement of Compliance:*

3. In progress: We are in the implementation phase.

*Self-assessment statement:*

DANS has a provenance document in which the way of publishing data at DANS is described in general. It also contains information on the way DANS deals with mutation and addition of (meta) data. The short document for the public is available at the website:

<http://www.dans.knaw.nl/sites/default/files/file/EASY/Provenance%20en%20dataverwerking%20DEF.pdf> (in Dutch). The extended and elaborate version for internal use and for colleague archives is available on request. DANS and its predecessors have been storing digital and analogue data for almost fifty years now. A provenance document for the period 1964-2005 is being prepared at this moment.

When a dataset is deposited, there are several steps that have to be taken by an archivist before the dataset can be published. The dataset is thoroughly checked for completeness, understandability and privacy-sensitive information by an archivist, both the data itself (file format, anonymisation of the file) and the accompanying metadata and documentation files. All data files uploaded by the depositor are always kept with the dataset as they were submitted originally, within a (virtual) folder named 'original'.

There are several security measures in EASY to make sure the data is authentic. When a file needs to be adapted, archivists work with a copy of the data, and never in EASY directly. They download a copy of the original file(s), work offline, and upload the adapted file(s) separate from the folder 'original'. Specific alterations are marked with additions to the file name, such as version numbers, or the addition 'ANON' when it concerns an anonymised file.

When archaeological data are being converted, the archaeological data manager adds file-specific metadata to every file, derived from a file list submitted with the dataset. The file-specific metadata includes information on the original file name and the original software that was used to create the file. File-specific metadata can be displayed with the files in EASY ('view details') and will be added to download packages.

Apart from the file name, EASY itself shows whether the upload has been done by an Archivist or the Depositor. This information is available to DANS archivists only, but can be shared with (re-) users of the data on request.

One way to ensure the data is treated properly is the obligatory workflow under the administration tab of each dataset. After processing a dataset, the archivist is required to mark his actions in a list of 12 options with checkboxes and a comment box. Five workflow-steps are mandatory (check for completeness of the submitted files; check for accessibility/readability/preferred formats; check for privacy-sensitive information; check for completeness of the submitted metadata/modify if desired; ensure that only the files which need to be published are accessible to users). The system only allows the archivist to publish the dataset after all five mandatory workflow boxes are checked (this will enable the 'publish' button). By checking any of the boxes, the name of the archivist and the date of checking appear next to the corresponding workflow step.

When there is extra information to be provided, the archivists write this down in the text pane 'Remarks' below the ticked boxes. When an e-mail from the data producer (or "depositor") on the dataset needs to be kept, they upload it in the dataset and keep it as a non-visible, non-downloadable file, thus to maintain the background of changes to the files and / or dataset.

At this moment most of the process is manual and subject to agreements and 'good practice' instead of automation, but DANS is working towards a more restrictive, automated model.

In regard to the authenticity of the depositor, usually depositors use their institutional e-mail address to create an EASY account and deposit the data. Often depositors contact DANS first before they start depositing their data. It could be that somebody claims to be a researcher and deposits data but the risk of discovery is quite high.

Currently, file formats are being investigated to make sure obsolete file formats are detected and converted in time. DANS makes sure that the content of the old files is equal to that of the new files.

#### **Planned activities:**

- DANS will look into further guarantees regarding the identity of a depositor and regarding the identity of the creator indicated by the depositor (Note that "data producer" can indicate either of them).
  
- Translate the Dutch provenance document to inform non-Dutch data depositors and users.

- Design and develop an automatic system for relating the workflow steps of a dataset to the underlying procedures as detailed in the provenance document.

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

The "3" rating seems appropriate.

### **13. The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.**

*Minimum Required Statement of Compliance:*

3. In progress: We are in the implementation phase.

#### **Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

DANS follows the OAIS reference model (<http://public.ccsds.org/publications/archive/650x0m2.pdf>) across the archival process. There is considerable support for Ingest, Archival storage, Data Management and Access. The Preservation Policy under construction (mentioned in Guideline 6) relates DANS's policy to these functions.

Whereas data curation, for example, is mostly done manually (see Guideline 12), other processes, like making datasets findable, guarding restrictions and distributing the data, are automated. Access to datasets is facilitated by the use of persistent identifiers (see Guideline 10). A content model has been created for collections, folders, and files.

EASY development is an ongoing process and also relates to developments outside DANS. DANS has a long-term development plan for EASY, which is the basis for year plans. The long-term development plan is available on request.

#### **Planned activities:**

Add provenance metadata to the existing descriptive metadata.

#### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

## 14. The data consumer complies with access regulations set by the data repository.

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

### Applicant Entry

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

#### **Metadata**

Metadata, as published in EASY, are always freely available for everyone without any registration. Metadata is the content of all fields under the 'Description' tab in every dataset in EASY.

#### **Data**

By registering, users agree to the DANS General Conditions of Use <http://www.dans.knaw.nl/en/content/dans-conditions-use-reuse-deposited-data>. Only registered data users are permitted to download datasets. There are two access categories for datasets in DANS EASY:

I. Open access: all registered data users have access to the dataset(s).

II. Restricted access: Only registered data users have access to the dataset(s) when this has been granted either by the holder of rights to the dataset(s) or by the data archive itself. The latter applies only to the 'archaeology group'. Access to this group is granted only to archaeology students at Dutch research universities or universities of applied sciences and to professional archaeologists working in the Netherlands.

The DANS General Conditions of Use are based upon the rules that are generally accepted with regard to the use of research data in scientific and scholarly research (see Guidelines 5 and 15).

For some datasets there are additional, stricter, conditions of use in force. These have been agreed upon with a number of depositors. These datasets mostly concern personal data of a sensitive nature.

### **Non-compliance with the Conditions of Use**

In the case of non-compliance with one of the DANS General Conditions of Use, the use of the dataset must be terminated immediately at the initial demand by DANS, according to article 8 of these conditions. DANS reserves the right, in such an event, to inform the user's employer. In the event of improper use of personal data, DANS also has the right to inform the Dutch Data Protection Authority (College Bescherming Persoonsgegevens).

### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

Having the statement about the measures undertaken for noncompliance with conditions of use strengthens the evidence here.

**15. The data consumer conforms to and agrees with any codes of conduct that are generally accepted in the relevant sector for the exchange and proper use of knowledge and information.**

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

**Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

- The DANS General Conditions of Use (see Guidelines 14 and 16 for references) are based upon the relevant laws of the EU and the Netherlands (Copyright, Database rights, Personal Data Protection Law) – see Guideline 5 for more information;
- The VSNU (Association of Universities in the Netherlands) Codes of Conduct: the [Nederlandse Gedragscode Wetenschapsbeoefening VSNU](#) - The Netherlands Code of Conduct for Scientific Practice and the [Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek](#) (in Dutch) - the Code of Practice for the use of personal data in scientific and scholarly research;
- The Open Access Principles (Berlin Declaration).

The VSNU “Code of Practice for the use of personal data in scientific and scholarly research” is a code of conduct which is derived from the Personal Data Protection Law of the Netherlands, specifically aimed at the processing of personal data for research purposes. DANS has Privacy Regulations <http://www.dans.knaw.nl/en/content/dans-privacy-regulations> (seen 29 October 2013) which give persons who have objections to the processing of his/her personal data by DANS the possibility to submit these objections to DANS. DANS shall subsequently inform the concerned party within four weeks whether the personal data, in accordance with the objection, are amended, supplemented or are protected or removed.

There is a special section at the website of DANS on how users should deal with personal data, as well as personal data contained in audio and visual data: <http://www.dans.knaw.nl/en/content/data-archive/legal-information> (seen 29 October 2013).

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*



## **16. The data consumer respects the applicable licences of the data repository regarding the use of the data.**

*Minimum Required Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

### **Applicant Entry**

*Statement of Compliance:*

4. Implemented: This guideline has been fully implemented for the needs of our repository.

*Self-assessment statement:*

For downloading data registration is obligatory (see Guideline 14). By registering, users agree to the DANS General Conditions of Use <http://www.dans.knaw.nl/en/content/dans-conditions-use-reuse-deposited-data> (seen 29 October 2013).

Several articles in the DANS General Conditions of Use contain conditions or instructions for the users how to make use of the data:

#### **4. Distribution of the dataset**

This article says that the dataset(s) may not be further distributed or made public without the prior written consent of the depositing party.

#### **5. Intended use of the dataset**

According to this article the dataset(s) may not be (re)sold or used for commercial purposes.

## 6. Personal data protection

This article stipulates that dataset(s) that contain personal information as referred to in the Personal Data Protection Act (Wet Bescherming Persoonsgegevens) may be used only for historical, statistical or scientific research (see Guideline 5). Persons who use datasets containing personal data are required to comply with the Code of Practice for the use of personal data in scientific and scholarly research (Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek) published by the VSNU (see Guideline 15 for references). The user undertakes to maintain confidentiality of all personal data that he/she processes. Additional requirements may be set regarding the use of datasets that contain special personal data, such as amongst others, data concerning religion, health or race.

For some datasets users have to comply with additional, stricter, conditions of use in force. These additional conditions are checked by the depositors themselves.

### Non-compliance with the Conditions of Use

In the case of non-compliance with one of the DANS General Conditions of Use, the use of the dataset must be terminated immediately at the initial demand by DANS, according to article 8 of these conditions.

DANS reserves the right, in such an event, to inform the user's employer. In the event of improper use of personal data, DANS also has the right to inform the Dutch Data Protection Authority (College Bescherming Persoonsgegevens).

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

*Comments:*