



Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository Huygens ING: eLaborate 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 April 29, 2014.

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:	Huygens ING: eLaborate
Seal Acquiry Date:	Apr. 29, 2014
For the latest version of the awarded DSA for this repository please visit our website:	http://assessment.datasealofapproval.org/seals/
Previously Acquired Seals:	None
This repository is owned by:	Huygens ING
	Netherlands
	T 070 – 331 5800
	F 070 – 385 4098
	E nfo@huygens.knaw.nl
	W http://www.huygens.knaw.nl/

Assessment

0. Repository Context

Applicant Entry

Self-assessment statement:

Huygens ING[1] is an institute of the Royal Academy of Arts and Sciences (KNAW)[2]. Its current mission is to investigate texts and sources from the past, using new methods and techniques. Huygens ING is committed to innovation in research methodology, but also to a better understanding of the culture and history of the Netherlands among a wide audience. The institute therefore wishes to conduct innovative and excellent research, and to reach out to society by stimulating innovation and making its products available to the public at large. It does not only deliver research value, but also social and cultural value.

Huygens ING combines traditional (and ever-rarer) knowledge in the humanities with expertise in the areas of information science and software development. The institute regards itself as a research laboratory in the humanities, and a space for experimentation.

Huygens ING was founded as the result of a merge in 2011 between ING – Institute for the History of the Netherlands, founded in 1902 – and the Huygens Institute for Dutch Literature and Intellectual History, founded in 1992. In over a century the institute and its predecessors have assembled a large collection of research data on many fields, ranging from medieval Dutch literary texts to data on Dutch-Indonesian colonial relations; from the place of Dutch feminist literature in a European context to the letter collections of Dutch 17th century scientists; and from the letters of Van Gogh to the Erasmus Opera Omnia. Although the institute considers itself primarily a research facility and does not maintain a public collection profile, Huygens ING actively maintains almost 200 digitally available resource collections. A full overview can be found at <http://www.huygens.knaw.nl/tools-en-data/>. The institute does not maintain physical collections.

For its research data, Huygens ING maintains two repositories. First, for storage of structured data like Persons, Places, Keywords and many other data records the institute has developed Timbuctoo[3]. Second, text editions, scans, transcriptions and annotations are stored in the online work environment eLaborate4.

As a prospective CLARIN center we adhere to the list of accepted formats and standards proposed by the CLARIN infrastructure[4] and by DANS.[5] Huygens ING also complies with the 'Open Access' policies stipulated by the KNAW.[6] Huygens ING has implemented the KNAW policies in its Data Policy Statement.[7] This document formulates the procedures for storage, publication and access to research data. Data deposits and the accompanying documentation in both repositories are evaluated by the coordinator research data collections (the data coordinator for short).

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

For long term digital preservation and archiving Huygens ING uses the services of DANS. Huygens ING and DANS are constructing a digital connection to archive individual datasets automatically. For each individual dataset the two parties have agreed on a license[8] for which an English version is also available.[9] DANS has a DSA certificate.[10]

[1] www.huygens.knaw.nl

[2] www.knaw.nl

[3] <http://repository.huygens.knaw.nl>

[4] <http://trac.clarin.nl/wiki/WikiStart#Formatsandstandards>

[5] <http://www.dans.knaw.nl/sites/default/files/file/EASY/DANS%20preferred%20formats%20NL%20DEF.pdf>

[6] <https://www.knaw.nl/en/topics/open-access-and-digital-preservation/open-access/policy>

[7] <http://www.huygens.knaw.nl/wp-content/uploads/2013/10/20120904-datanotitie-Huygens-ING.pdf>

[8] http://www.dans.knaw.nl/sites/default/files/file/EASY/Licentieovereenkomst%20DANS_NL_4_2.pdf

[9] <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>

[10] https://assessment.datasealofapproval.org/assessment_101/seal/pdf/

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

This is a very useful statement setting the context for the repository and the assessment.

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:

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

Self-assessment statement:

Since Huygens ING does not have a public collection profile, storage of datasets in its repositories are part of the institute's ongoing research activities. All research project plans proposed by scholars contain documentation on data collection and/or curation of existing sources. This paragraph is always written in combination with data specialists from IT. Due to the innovative nature of research, new or changing standards and/or archaic systems that contain the data, Huygens ING has consciously not set any requirements for deposit packages. These details will be included in the data paragraph by our data specialists. This way we make sure that the proposal contains enough funding for the required curation. For each requested deposit the data coordinator evaluates the documentation on the data collecting practices and used standards. When and where necessary the data coordinator will request more information or consults internal or external specialists to assess the content of the data.

For his assessment the data coordinator follows general guidelines and standing best practices in the field. Huygens ING projects are usually highly innovative which means that they often set data standards instead of following them.

The repositories do not (and cannot) systematically verify whether the data is collected according to these guidelines.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

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

Self-assessment statement:

As a prospective CLARIN center Huygens ING adheres to the list of accepted formats and standards proposed by the CLARIN infrastructure^[1] - CLARIN maintains a short guide on these standards.^[2] All research project plans proposed by scholars contain documentation on the standards used in the project. This paragraph is always written in combination with data specialists from IT. Due to the innovative nature of research, new or changing standards and/or archaic systems that contain the data, Huygens ING may allow other standards as far as they are useful for literary and historical academic research. In these situations prior approval of the data coordinator is required.

During ingestion the format and validity of the ingested resources is checked. Huygens ING is in the process of automating this check and ingestion process. When necessary it is performed manually by a data specialist. Some scholars have acquired existing – usually old – data sources. Issues with these file formats are identified in the research project plan. The data coordinator is involved with calculating the required funding to allow for curation of the data to accepted formats. Given enough funding, Huygens ING in principle will not exclude any file format for ingestion.

^[1] <http://www.clarin.eu/sites/default/files/Standards%20for%20LRT-v6.pdf>

^[2] <http://www.clarin.eu/sites/default/files/standards-CLARIN-ShortGuide.pdf>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

The data producer is required to provide metadata. Metadata descriptions are generally created for bundles of resources that belong together – e.g. text editions containing scans and transcriptions. The eLaborate repository is designed for the creation of metadata descriptions. Dublin Core is the minimal requirement^[1] and Huygens ING follows as much as possible the specification of Qualified Dublin Core.^[2] To help researchers generate the DC metadata, the eLaborate editing environment provides data producers with a form to register the required project data. Text editions cannot be published while this form is empty. The data coordinator evaluates the input with each deposit and the accompanying documentation.

As a prospective CLARIN center Huygens ING automatically converts the metadata of CLARIN funded projects in eLaborate and Timbuctoo to CMDI and exposes the results through OAI-PMH.

[1] http://elaborate.huygens.knaw.nl/?page_id=107

[2] <http://dublincore.org/documents/dcmi-terms>

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:

Huygens ING^[1] is an institute of the Royal Academy of Arts and Sciences (KNAW).^[2] Its current mission is to investigate texts and sources from the past, using new methods and techniques. Huygens ING is committed to innovation in research methodology, but also to a better understanding of the culture and history of the Netherlands among a wide audience. The institute therefore wishes to conduct innovative and excellent research, and to reach out to society by stimulating innovation and making its products available to the public at large. It does not only deliver research value, but also social and cultural value.

Although the institute considers itself primarily a research facility and does not maintain a public collection profile, Huygens ING actively maintains almost 200 digitally available resource collections. A full overview can be found at <http://www.huygens.knaw.nl/tools-en-data/>. The institute does not maintain physical collections.

For long term digital preservation and archiving Huygens ING uses the services of DANS. Huygens ING and DANS are constructing a digital connection to archive individual datasets automatically. For each individual dataset the two parties have agreed on a license^[3] for which an English version is also available.^[4] DANS has a DSA certificate.^[5]

^[1] www.huygens.knaw.nl

[2] www.knaw.nl

[3] http://www.dans.knaw.nl/sites/default/files/file/EASY/Licentieovereenkomst%20DANS_NL_4_2.pdf

[4] <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>

[5] https://assessment.datasealofapproval.org/assessment_101/seal/pdf/

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

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:

Both eLaborate and Timbuctoo are the property of the Huygens Institute for the History of the Netherlands (Huygens Instituut voor Nederlandse Geschiedenis), which in its turn is a research institute of the Royal Academy of Arts and Sciences (Koninklijke Nederlandse Akademie voor Wetenschappen). The repositories are primarily funded by Huygens ING. Both internal and external deposits in the repositories are made through research project proposals. The proposals are evaluated by the data coordinator and accepted or rejected by Huygens ING management. All legal agreements with the depositor – compliance with national laws, ethical issues etc. – are part of this proposal. Due to the highly varied nature of innovative research proposals, Huygens ING does not use standard legal agreements with depositors.

When applicable research proposals contain a paragraph on the Code of Conduct for Use of Personal Data in Scientific Research, which is available in Dutch as *Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek VSNU*.^[1]

Huygens ING complies with the ‘Open Access’ policies as stipulated by the KNAW. The institute allows depositors to restrict access to data at various levels.

[1] <http://www.vsnu.nl/code-pers-gegevens.html>

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:

Huygens ING has implemented KNAW policies in its Data Policy Statement^[1]. This document formulates the procedures for storage, publication and access to research data. Physical storage of data in both repositories is provided through the KNAW information and automation service I&A. KNAW I&A uses the services of Vancis^[2]. Vancis offers data services for universities, companies and educational institutions. Backups are made every day – during the night – and access from outside to the servers for maintenance is restricted. Due to the sensitive nature of data storage procedures, KNAW policy dictates that they remain confidential.

For long term digital preservation and archiving Huygens ING uses the services of DANS. Huygens ING and DANS are constructing a digital connection to archive individual datasets automatically. For each individual dataset the two parties have agreed on a license^[3] for which an English version is also available^[4]. DANS has a DSA certificate^[5].

[1] <http://www.huygens.knaw.nl/wp-content/uploads/2013/10/20120904-datanotitie-Huygens-ING.pdf>

[2] <http://www.vancis.nl>

[3] http://www.dans.knaw.nl/sites/default/files/file/EASY/Licentieovereenkomst%20DANS_NL_4_2.pdf

[4] <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>

[5] https://assessment.datasealofapproval.org/assessment_101/seal/pdf/

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

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:

For long term digital preservation and archiving Huygens ING uses the services of DANS. Huygens ING and DANS are constructing a digital connection to archive individual datasets automatically. For each individual dataset the two parties have agreed on a license^[1] for which an English version is also available^[2]. DANS has a DSA certificate^[3]. In its DSA DANS explains the processes and procedures for (1) managing data storage and (2) the long term preservation of the digital assets, including implementing functions for recognition of file formats and for converting files that are in formats which are in danger of becoming obsolete.

Currently, Huygens ING data specialists transfer individual datasets manually to DANS and licenses are agreed on a case-by-case basis. Huygens ING and DANS are in the process of constructing a digital connection to archive datasets automatically. A blanket agreement for all these datasets is the subject of discussions.

[1] http://www.dans.knaw.nl/sites/default/files/file/EASY/Licentieovereenkomst%20DANS_NL_4_2.pdf

[2] <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>

[3] https://assessment.datasealofapproval.org/assessment_101/seal/pdf/

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

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

Self-assessment statement:

There is a common generic workflow, however, due to the highly varied nature of innovative research at Huygens ING the workflow is not registered in a general procedure. Instead a description of the data life cycle is a required part of research project proposals. The data coordinator evaluates the proposed data life cycle before the plan is accepted or rejected by Huygens ING management.

The depositor determines who can access the material and - where applicable - is also responsible for protecting the privacy of any subjects appearing in the texts.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

Huygens ING complies with the 'Open Access' policies as stipulated by the KNAW. The institute allows depositors to restrict access to data at various levels. All distributed copies are stored under the agreement that they are subject to the same restrictions. Access restrictions are part of the research project proposal.

For long term digital preservation and archiving Huygens ING uses the services of DANS. Huygens ING and DANS are constructing a digital connection to archive individual datasets automatically. For each individual dataset the two parties have agreed on a license^[1] for which an English version is also available.^[2]

Huygens ING strives for 24/7 availability for both the eLaborate and Timbuctoo repositories. In emergency situations Huygens ING depends on KNAW I&A to restore physical access. Due to its sensitive nature the KNAW I&A crisis management plan is considered confidential. Restoration of back-ups is the responsibility of Huygens ING. In case of full system breakdown Huygens ING expects to be back online within several days.

[1] http://www.dans.knaw.nl/sites/default/files/file/EASY/Licentieovereenkomst%20DANS_NL_4_2.pdf

[2] <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>

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:

Both the eLaborate and Timbuctoo repositories provide various ways of utilizing the research data via online tools as well as by downloading the data in formats commonly used by research communities. Both repositories provide a RESTfull web api which exposes data in both XML and JSON. Most research proposals contain a web research environment also allowing human access to the data with faceted and full text search features.

Unique persistent identifiers according to the Handle system are provided for each data object.

As a prospective CLARIN center, Huygens ING exposes CMDI metadata through OAI-PMH for all CLARIN funded projects.

For long term digital preservation and archiving Huygens ING uses the services of DANS. For each individual dataset the two parties have agreed on a license. The preserved archived version of the dataset is accessible through the DANS EASY system.

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:

Once deposited data objects are never changed. All editions of a file are automatically stored as new versions - each with a unique persistent identifier. Old versions remain fully available.

Huygens ING has no current plans to provide MD5 checksums to guarantee data integrity but will consider this in the future.

The availability of all servers is monitored continuously.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

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

Self-assessment statement:

All deposited data is subject to acceptance by the data coordinator. Authenticity of the data is checked as part of the process. Most research project proposals contain a web research environment allowing the creation of new data, updates and deletions by the depositor or (a group of) scholar(s) to who the depositor has delegated crud-permissions. Changes are recorded in different versions containing basic time, user- and system identification metadata. Each version receives a persistent identifier.

As a prospective CLARIN center, Huygens ING identifies unique users through Federated Identity Management.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

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

Self-assessment statement:

Neither eLaborate nor Timbuctoo was designed and implemented as an archive. The architecture of both systems does not explicitly follow OAIS or other archival standards.

For archiving purposes and the implementation of internationally accepted archival standards, Huygens ING uses the services of DANS. For each individual dataset the two parties have agreed on a license^[1] for which an English version is also available.^[2] DANS has a DSA certificate.^[3]

Currently, Huygens ING data specialists transfer individual datasets manually to DANS. Huygens ING and DANS are in the process of constructing a digital connection to archive datasets automatically. Datasets stored at DANS do comply with the required standards.

[1] http://www.dans.knaw.nl/sites/default/files/file/EASY/Licentieovereenkomst%20DANS_NL_4_2.pdf

[2] <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>

[3] https://assessment.datasealofapproval.org/assessment_101/seal/pdf/

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

Most of the data in the eLaborate repository is protected; an account is necessary to get access to the data. User identification is guaranteed through federated identity management. For Timbuctoo most data in the repository is open access, no login is required to read the data and earlier versions.

All datasets have a Creative Commons license applied to them. Huygens ING does not actively carry out any checks if users comply, but in the event of demonstrable abuse, Huygens ING will first contact the user personally but will eventually refer to the user's employer. In the event of a violation of the Personal Data Protection Act, Huygens ING will report abuse to the employer and the Data Protection Authority.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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:

Huygens ING complies with the 'Open Access' policies stipulated by the KNAW[1]. Huygens ING has implemented the KNAW policies in its Data Policy Statement[2].

Huygens ING also complies to the Code of Conduct for Use of Personal Data in Scientific Research, Which is available in Dutch as *Gedragcode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek VSNU*[3]. For both repositories Huygens ING makes use of Creative Commons licenses.

[1] <https://www.knaw.nl/en/topics/open-access-and-digital-preservation/open-access/policy>

[2] <http://www.huygens.knaw.nl/wp-content/uploads/2013/10/20120904-datanotitie-Huygens-ING.pdf>

[3] <http://www.vsnu.nl/code-pers-gegevens.html>

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:

Where applicable, the data consumer is made aware of usage restrictions for the data she/he has gotten access to. Web research environments provided by Huygens ING and accessing either – or both – repositories, contain a page informing the user of the applicable licenses and general terms and conditions of use. Huygens ING does not actively check user compliancy with these agreements. In the case of demonstrable abuse Huygens ING will report the user to his or her employer. In the event of a violation of the Personal Data Protection Act, Huygens ING will report abuse also to the Data Protection Authority

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments: