



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

The Data Seal of Approval board hereby confirms that the Trusted Digital repository TRAILS 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 May 21, 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:   | <a href="#">DSA-booklet_2014-2017.pdf</a>   |
| All Guidelines Documentation:   | <a href="#">Documentation</a>   |
| Repository:   | TRAILS  |
| Seal Acquiry Date:  | May. 21, 2014   |
| For the latest version of the awarded DSA for this repository please visit our website: | <a href="http://assessment.datasealofapproval.org/seals/">http://assessment.datasealofapproval.org/seals/</a>   |
| Previously Acquired Seals:  | None  |
| This repository is owned by:  | <b>TRAILS</b> <ul style="list-style-type: none"><li>• TRAILS<br/>Sector C- CC 33 PO Box 30.001<br/>9700 RB Groningen<br/>The Netherlands</li></ul><br>T +31 50 3611205<br>E <a href="mailto:trails@umcg.nl">trails@umcg.nl</a><br>W <a href="http://www.trails.nl/">http://www.trails.nl/</a> |

# Assessment

## 0. Repository Context

### Applicant Entry

*Self-assessment statement:*

TRAILS (Tracking Adolescents' Individual Lives Survey) is an ongoing, multidisciplinary research on the psychological, social and physical development of children towards adulthood. TRAILS started in 2001, and is a cooperation between University Medical Center Groningen and the Erasmus Medical Center Rotterdam. For an overview of all institutions involved in TRAILS please refer to the TRAILS website (<http://www.trails.nl/en/hoofdmenu/over-trails/trails-team/betrokken-instellingen>). The TRAILS infrastructure is situated in Groningen. In terms of the Open Archival Information System (OAIS) model, TRAILS fulfills the functional roles of data production, data management and data consumption.

Currently, the data available to external data consumers are archived in DANS (Data Archiving and Networked Services) EASY (Electronic Archiving System) (see, for instance: <http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-o7g-lqz>). The data are only available through TRAILS, however. The appropriate procedures for data access are available from the TRAILS website (see: <http://www.trails.nl/en/hoofdmenu/trails-results/data-use>). When its infrastructure is terminated, TRAILS will outsource the data management function (long-term archiving and dissemination of TRAILS data) to DANS EASY, for which TRAILS has a data agreement with DANS.

Outsource partner DANS is involved with the development of the Digital Seal of Approval. Its online archiving system EASY was awarded the Digital Seal of Approval 2014-2015 in November 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)

**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:*

The organization of the TRacking Adolescents' Individual Lives Survey (TRAILS) is such that the functional roles of data production and data archive are to a large extent concentrated within the same organizational role. TRAILS does not archive data from external data producers.

TRAILS collects data once every 2 to 3 years by means of self-administered questionnaires. Measurement waves are usually complemented with additional questionnaires, interviews, experiments, or psychological tests. Which instruments to include in a questionnaire is determined by the TRAILS management team, in close consultation with the TRAILS domain holders. Whenever possible, validated measurement instruments that are well-established in the scientific discipline of interest are used in data collection. Prior to data collection domain holders have to provide the data producers and data repository with a Dutch version of the instrument. In addition, additional information has to be provided, such as references (e.g. validation, translation), instructions for programming (e.g. lay-out, routing), and instructions for data processing (e.g. coding, scale construction). During and after data collection, data producers provide the data repository with data files, and detailed information regarding (non)response, applicable protocols, evaluation reports, and any other relevant documentation. All potentially relevant information is processed and checked by the data repository, and made available to data consumers either by including it in the relevant codebook, or by noting its availability. These procedures have been in use for many years, and have proven to be effective. An elaborate data management protocol based on the Open Archival Information System (OAIS) model (see guideline 13) in which these and other procedures are formally laid out, is currently in preparation and will be completed by September 2014.

Furthermore, the data producers, data repository and (internal) data consumers regularly cooperate to publish articles in which the scientific and scholarly quality of the TRAILS data are discussed. De Winter et al. (2005) discuss sampling strategy and non-response of the first measurement wave. They conclude that extended

recruitment efforts, which raised the response rate from 66% to 76%, resulted in the recruitment of more vulnerable children in the study. Huisman et al. (2008) give a general description of the first two measurement waves of the TRAILS study. Nederhof et al. (2012) build upon De Winter et al. (2005), and discuss participant retention and attrition at the time of the fourth measurement wave. They report that the positive effects of the extended recruitment efforts at the first measurement wave still existed eight years later, and that extended recruitment efforts at the fourth measurement wave had a similar effect to those at the first measurement wave. Finally, Ormel et al. (2012) provide a general overview of the TRAILS study up to and including the fourth measurement wave, and present an integrated discussion of study findings from the first 10 years of TRAILS. They conclude that TRAILS has contributed substantially to the understanding of mental health problems and social development during adolescence.

Because the main contributors of TRAILS are (university) medical centers, TRAILS is subject to much stricter ethical regulations than, for example, studies performed within the discipline of the social sciences. Prior to each measurement wave, permission needs to be granted by the Central Committee on Research Involving Human Subjects (CCMO). The CCMO protects subjects taking part in medical research by reviewing the research on the basis of the statutory provisions laid down for them and taking account of the interests of medical progress (examples of information reviewed by the CCMO are research protocols, consent forms, questionnaires and participant information). The CCMO sees to it that medical research is in compliance with national law (among which the Medical Research Human Subjects Act) and international law.

The Trusted Digital Repository (TDR) Data Archiving and Networked Services (DANS) Electronic Archiving System (EASY) (see: <https://easy.dans.knaw.nl/ui/home>) will provide the long-term preservation and dissemination of the data (DANS EASY has received the DSA in April 2011). When depositing data, TRAILS will conform to the requirements set by DANS EASY as indicated in the license agreement (see: <http://www.dans.knaw.nl/en/content/dans-licence-agreement-deposited-data>) and the data depositing instructions (see: <http://www.dans.knaw.nl/en/content/data-archive/depositing-data>). TRAILS has already deposited a substantial part of its archive with DANS EASY as part of its dissemination strategy for data consumers external to the TRAILS consortium, although TRAILS will continue to distribute data from its own archive until the study ends.

Full references to the key TRAILS publications:

- De Winter, A.F., Oldehinkel, A.J., Veenstra, R., Brunnekreef, J.A., Verhulst, F.C., & J. Ormel (2005). Evaluation of non-response bias in mental health determinants and outcomes in a large sample of pre-adolescents. *European Journal of Epidemiology*, 20(2), 173-181. doi: 10.1007/s10654-004-4948-6
- Huisman, M., Oldehinkel, A.J., Winter, A. de, Minderaa, R.B., Bildt, A. de, Huizink, A.C., Verhulst, F.C., & J. Ormel (2008). Cohort profile: The Dutch 'TRacking Adolescents' Individual Lives' Survey'; TRAILS. *International Journal of Epidemiology*, 37(6), 1227-1235. doi: 10.1093/ije/dym273
- Nederhof E., Jorg F., Raven D., Veenstra R., Verhulst F.C., Ormel J., & A.J. Oldehinkel (2012). Benefits of extensive recruitment effort persist during follow-ups and are consistent across age group and survey method. The TRAILS study. *BMC Medical Research Methodology*, 12:93. doi:10.1186/1471-2288-12-93
- Ormel, J., Oldehinkel, A. J., Sijtsema, J. J., Raven, D., Van Oort, F. V. A., Veenstra, R., Vollebergh, W. A. M., & Verhulst, F. C. (2012). The TRacking Adolescents' Individual Lives Survey (TRAILS). Design, current status, and selected findings. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51, 1020-1036.

## 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)

## **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:*

All information is to be provided to the data repository in a preferred format (mostly SPSS SAV, MS Word, or MS Access) or another format that is easily converted (such as plain text, CSV, MS Access, RTF, or PDF).

When depositing data, TRAILS will supply data and documentation in a preferred or acceptable format as requested by DANS EASY (most likely SPSS portable for data and MS Word for documentation; see: <http://www.dans.knaw.nl/en/content/data-archive/depositing-data>).

### **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:*

Prior to data collection domain holders have to provide the data producers and data repository with a Dutch version of the instrument. In addition, additional information has to be provided, such as references (e.g. validation, translation), instructions for programming (e.g. lay-out, routing), and instructions for data processing (e.g. coding, scale construction). During the data collection process, TRAILS staff are required to fill out contact details on a Contact Status Form (CSF). Key information such as changes in contact details, appointments, and reasons for non-response are also entered into the central logistics database maintained at the TRAILS secretary. During and after data collection, data producers provide the data repository with data files and detailed information regarding (non)response, applicable protocols, evaluation reports, and any other relevant documentation. The data repository provides the data producers with codebook guidelines. All documentation that is supplied by the data producers is checked by the data repository, and, if necessary, additional information or further clarification is requested by the data repository. These procedures have been in use for many years, and have proven to be effective. An elaborate data management protocol based on the Open Archival Information System (OAIS) model (see guideline 13) in which these and other procedures are formally laid out, is currently in preparation and will be completed by September 2014.

When depositing data, TRAILS will supply meta data according to the requirements set by DANS EASY. These requirements are based on the Qualified Dublin Core (see: <http://dublincore.org/documents/dcmi-terms/>).

#### **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)



#### **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:*

One of the main goals of TRAILS is to contribute to the knowledge of the development of emotional and behavioral problems and the (social) functioning of preadolescents into adulthood, their determinants, and underlying mechanisms. In order to be able to complete this goal, TRAILS is financially supported by various grants from multiple (scientific) institutions. Next to, for example, publishing research findings in scientific journals, TRAILS therefore perceives properly documenting, archiving, and providing access to the collected data as a duty towards the supporting institutions in particular and the scientific community and society in general. This duty was translated into another of the main objectives of TRAILS: “advancing knowledge accumulation and knowledge dissemination”.

The long-term preservation and dissemination of research data will be outsourced to DANS EASY, which has an explicit mission statement regarding digital archiving.

#### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

## **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:*

TRAILS is not a legal entity. The main participants in the TRAILS consortium are departments of the University Medical Center Groningen (UMCG) and the Erasmus Medical Center, which are both legal entities.

As explained in the answer to guideline 1, the CCMO monitors the compliance of TRAILS with national and international law. TRAILS staff are required to sign a non-disclosure agreement. End users are required to agree with the TRAILS general conditions (see: <http://www.trails.nl/en/hoofdmenu/trails-results/data-use>). The general conditions cover the access to TRAILS data, demands placed on the manuscript in which TRAILS data are reported, and codes of conduct regarding the use of TRAILS data. In case of non-compliance, TRAILS will call upon the data users to immediately cease the use of TRAILS data, and destroy any copies of the TRAILS data. TRAILS reserves the right to inform the data users' employer and, in case of Dutch Data Protection Act infringements, the Dutch Data Protection Authority.

TRAILS will outsource long-term preservation and dissemination of research data to DANS EASY when its infrastructure (including data management) is terminated. When depositing data, TRAILS will conform to the requirements set by DANS EASY. Data users who want to use TRAILS data will need to comply with the terms and conditions set by DANS EASY, in addition to those set by TRAILS.

## **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:*

## **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:*

All files obtained from data producers are stored in the archive in their original condition. After any further data validation procedures the syntax, processed data, and documentation files are also archived. File names always start with the date of creation. A current copy of each processed data file is stored separately within the archive, and can be used for data and information requests. Access to the TRAILS archive is restricted to members of the data management group.

The TRAILS archive is stored on the central IT infrastructure of the University of Groningen, which is located on two separate sites. An automated back-up is performed on a daily basis.

These procedures have been in use for many years, and have proven to be effective. An elaborate data management protocol based on the Open Archival Information System (OAIS) model (see guideline 13) in which these and other procedures are formally laid out, is currently in preparation and will be completed by September 2014.

### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

This guideline is accepted on the assumption that such documented procedures are being prepared and will be made available.

**Data Seal of Approval Board**

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

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

TRAILS has a data agreement with DANS in which the archiving and dissemination of TRAILS data through DANS EASY is recorded. In general, each time a new assessment wave of TRAILS data becomes available to researchers from within the TRAILS consortium, data and documentation from the previous assessment wave are uploaded to DANS EASY (see, for instance: <http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-o7g-lqz>). As long as the TRAILS infrastructure exists, DANS EASY is used for archiving previous assessment waves and for communicating the availability of these data to external researchers. TRAILS data are not publicly available through DANS EASY, but can be requested by following the procedure as outlined in the TRAILS general conditions (see: <http://www.trails.nl/en/hoofdmenu/trails-results/data-use>). When the TRAILS study ends and its data infrastructure is terminated, DANS EASY will take over the responsibilities of archiving and dissemination.

Recently, demands placed on research projects such as TRAILS have changed in order to, among other things, better insure the adequate documentation, archiving and dissemination of data. TRAILS and DANS are therefore in the process of upgrading the current data agreement into a data contract. Preparing a formal data management protocol and applying for a Data Seal of Approval are part of this process.

### **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:*

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

*Self-assessment statement:*

The workflows across the data life cycle are described in multiple, successive internal protocols. First, for each assessment wave the general logistical procedures are described in a research protocol. This protocol describes the ways in which respondents are approached, the type of data to be collected, and the instruments used to collect the data. Once the research protocol is approved by the CCMO (see guideline 1), its contents are translated into protocols in which detailed procedures are specified.

The data life cycle is captured in four separate (sets of) protocols.

First, applicable prior to data collection, instrument protocols prescribe how the various instruments should be programmed (such as online questionnaires) and/or installed (such as a computer assisted personal interview [CAPI]).

Second, contact protocols prescribe when to contact participants, how to handle when contacting participants, how to document contact (attempts) with participants in the Contact Status Form (see guideline 3), and how to process contact (attempts) and their results in the logistical database.

Third, a data processing protocol prescribes how the data that were collected are to be transferred from the data production function to the data archiving function, and how information regarding the process of data collection is to be transferred back to the data production function by the data archiving function.

Fourth, a set of data management protocols explains the structure of the network drive on which the data are placed. Furthermore, procedures regarding data archiving, data file naming (version management), data cleaning, and data dissemination are prescribed. Also, the agreements with DANS are established.

Next to these protocols, procedures of interest to researchers who want to use TRAILS data are explained in the TRAILS general conditions (also see guidelines 5, 14, 15, and 16). The TRAILS general conditions are freely available on the TRAILS website.

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

We would suggest that these internal protocols are either made publically available or the reasons why this is not practical are defined in future submissions for the DSA

## **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:*

During and after data collection, the data production function and data archiving function routinely exchange information and data (also see guidelines 3 and 8). Once data and other relevant information are transferred, the data archiving function assumes the responsibility for the archiving and processing of the data, for adequately documenting the data and any data collection and processing procedures, and for providing access to the data. The data archiving function is responsible for the integrity of the data and information it has received. If these are considered incomplete or unclear, the data archiving function actively seeks to retrieve additional data or information.

As long as the TRAILS infrastructure exists, TRAILS will be responsible for archiving and access. Data that are available for scientific research by external data users will be deposited in DANS EASY. Although DANS will receive a copy of the data, TRAILS is responsible for providing data access to external data users ('other access' option in DANS EASY). TRAILS will outsource long-term preservation and dissemination of research data to DANS EASY when its infrastructure (including data management) is terminated. TRAILS will however always remain the proprietor of its 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:*

Data are provided to end users in SPSS format. SPSS has the advantage of being one of the most popular statistical software packages used in the health and social sciences. Furthermore, SPSS allows for easy conversion into the formats of other statistical packages. Usually, end users convert the data into the format of their preference by themselves if they wish.

TRAILS does not offer the possibility for OAI harvesting and deep searching. On the TRAILS website a schematic overview of the available data is available. After registering, interested researchers can download the TRAILS codebooks from DANS EASY. Also, the data manager is able to give advice regarding variable selection.

End users cannot refer to specific TRAILS data files using persistent identifiers. Persistent identifiers are created after depositing data in DANS EASY, but these refer to the metadata and not to unique data files. Usually, end users refer to the TRAILS data in general by including references to the key TRAILS publications as mentioned in guideline 1. Where applicable, researchers are stimulated to refer to previous TRAILS research.

### **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)

## 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:*

Files with processed data are obtained by running data processing syntaxes on original data files. All file names start with the date of creation, and old files are never deleted (also see guideline 6). All obsolete files remain in the archive, but working directories only contain the most recent data files. Only data from the most recent files are distributed to end users. Once a manuscript is accepted for publication, researchers are obliged to send their syntax to the data manager to be archived. Although these procedures are aimed at allowing for the replication of any publication on TRAILS data, the responsibility for adequately archiving the files and information on which the publication is based remain with the corresponding author of the respective publication.

Currently, TRAILS does not employ structural checks of data integrity, such as the periodical verification of data through the use of checksums. TRAILS does, however, acknowledge the importance of such verification procedures, especially because the earliest data files in the archive are over 12 years old, and because the archive is still growing. TRAILS undertakes to extensively inquire into checksum verification in the second half of 2014, and to implement checksum verification in the first half of 2015. The aim is to have a checksum verification procedure operational before July 2015.

### **Reviewer Entry**

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

Accept

#### *Comments:*

Accepted due to the clear intentions regarding additional data integrity measures.

## 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:*

When the responsibility for the data files is transferred from the data production function to the data archiving function, the data manager checks the data integrity (see also guidelines 3 and 9). The TRAILS data manager is responsible for the TRAILS data base and any changes made to the data. Data users are requested to inform the data manager of any errors in the data they discover. Any data changes are recorded in a syntax file, and are noted in the relevant codebook. Obsolete versions of all files remain in the archive. Because the date of creation is part of the file name of any file, changes to the file can be backtracked through time (also see guideline 11). Updated data files are also added to DANS EASY (obsolete files are retained in DANS EASY as well). Data users cannot change data in the TRAILS data base themselves.

A number of measures are in place to minimize inadvertent side effects. First, only a very limited number of people have access to the data archive (the data management group, consisting of the TRAILS primary investigator, the data manager, the ICT manager, and the data collection coordinator). Second, working directories are maintained with only current versions of data files. The full data archive, including obsolete and current files, therefore is not used for regular activities (such as data dissemination) performed by the data archiving function. Finally, all files in the data archive are write-protected.

Because TRAILS collects its own data, and because data from other data producers is not allowed into the TRAILS archive, there is no need for TRAILS to check the identity of depositors. Also, TRAILS does not maintain links to metadata and other data sets.

### **Reviewer Entry**

*Accept or send back to applicant for modification:*

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

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

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

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:*

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

*Self-assessment statement:*

Although in the past TRAILS did not apply a standard like OAIS, the TRAILS structure resembles the OAIS to a large degree. The following is a short description of TRAILS in terms of the OAIS.

TRAILS in terms of OAIS

Because TRAILS collects its own data, the organization of the is such that the data production function and the data archiving function largely overlap. TRAILS is relatively small compared to data archives such as DANS EASY. Therefore, not only are the six functional components of the data archiving function concentrated within the same organizational role, but the data production function and the data archive function largely overlap in the same organizational role as well.

Broadly, two types of data consumers can be distinguished. Internal consumers are part of the TRAILS consortium. Data are collected according to the needs and demands of the TRAILS consortium (internal data consumers can therefore also be involved with some data production and data archiving functions). External consumers are consumers who wish to use the TRAILS data, but are not part of the TRAILS consortium.

Ingest

**Data Seal of Approval Board**

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

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

All information is to be provided in a recommended format (mostly SPSS SAV, MS Word, or MS Access) or another format that is easily converted (such as plain text, CSV, MS Access, RTF, or PDF). If any data cleaning was performed, raw data, syntax, processed data, and documentation files must be supplied. Subsequently, data files are validated to assure they are not corrupted and case identification is correct.

#### Archival Storage

All files obtained from data producers are stored in the archive in their original condition. After any further data validation procedures the syntax, processed data, and documentation files are also archived. A current copy of each processed data file is stored separately within the archive, and can be used for data and information requests. Obsolete files always remain in the archive. The archive is stored on the central IT infrastructure of the University of Groningen, which is located on two separate sites.

#### Data Management

If necessary, data files are cleaned according to data cleaning protocols. End user documentation is prepared based on the files supplied by data producers and data cleaning. Access queries are performed and information requests are fulfilled.

#### Preservation Planning

The TRAILS data will at all times remain the property of the main participants of the TRAILS consortium. TRAILS will continue to distribute data from its own archive until its infrastructure (including data management) is terminated. Upon termination, the TRAILS data are archived by the main participants of the TRAILS consortium for the long-term. In addition, long-term archiving of, and access to the TRAILS data will be outsourced to DANS EASY. TRAILS is in the process of updating its existing agreements with DANS.

## Access

TRAILS data are available for scientific researchers affiliated with TRAILS (internal consumers) and scientific researchers in the medical and social sciences in general (external consumers). The availability of data for external consumers is communicated through DANS EASY. All consumers are bound by the TRAILS General Conditions (see: <http://www.trails.nl/en/hoofdmenu/trails-results/data-use>), and must submit a publication proposal for approval by the TRAILS management team. Only the data specified in approved proposals are distributed by the data manager.

## Administration

The five functional components of the data archive mentioned above are coordinated by the administration function. It is crucial that the data production function and the data archive interact smoothly during data collection periods. Another important responsibility is the implementation of new policies in order to satisfy increased demands placed on the archive (such as to meet the requirements for a DSA).

TRAILS is currently preparing an elaborate data management protocol based on the Open Archival Information System (OAIS) model.

## 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:*



## 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:*

In order to gain access to TRAILS data, data users have to agree with the TRAILS general conditions (see: <http://www.trails.nl/en/hoofdmenu/trails-results/data-use>). The general conditions cover the access to TRAILS data, demands placed on the manuscript reporting TRAILS data, and codes of conduct regarding the use of TRAILS data. Upon submitting a publication proposal, the data user agrees with the TRAILS general conditions.

Researchers can only request TRAILS data from the TRAILS data manager after approval of the publication proposal by the TRAILS MT. All data users who fail to comply with the TRAILS access regulations will not receive approval of their publication plan. Any data request without such approval will be denied.

TRAILS will outsource long-term preservation and dissemination of research data to DANS EASY when its infrastructure (including data management) is terminated. When depositing data, TRAILS will conform to the requirements set by DANS EASY. Data users who want to use TRAILS data will need to comply with the terms and conditions set by DANS EASY, in addition to those set by TRAILS.

### 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)

**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:*

Upon submitting a publication proposal, the data user agrees with the TRAILS general conditions. The TRAILS general conditions include a section on codes of conduct regarding the use of TRAILS data. A specific reference is made to the “Code of Conduct for Scientific Practice” (see: <http://www.vsnu.nl/Universities/Quality-assurance/Code-of-conduct-for-scientific-practice-1.htm>) as published by the Association of Universities in the Netherlands.

In case the codes of conduct are not adhered to, the TRAILS MT’s approval of the data users’ publication plan is revoked immediately, thereby also withdrawing the data users’ right to use the TRAILS data. TRAILS will call upon the data user to immediately cease the use of TRAILS data, and destroy any copies of the TRAILS data. TRAILS reserves the right to also inform the data users’ employer and, in case of Dutch Data Protection Act infringements, the Dutch Data Protection Authority.

TRAILS will outsource long-term preservation and dissemination of research data to DANS EASY when its infrastructure (including data management) is terminated. When depositing data, TRAILS will conform to the requirements set by DANS EASY. Data users who want to use TRAILS data will need to comply with the terms and conditions set by DANS EASY, in addition to those set by TRAILS.

**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:*

## **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:*

Upon submitting a publication proposal, the data user agrees with the TRAILS general conditions. Please refer to points 14 and 15 for an elaboration of agreement with the TRAILS general conditions and actions undertaken in the event of non-compliance.

TRAILS will outsource long-term preservation and dissemination of research data to DANS EASY when its infrastructure (including data management) is terminated. When depositing data, TRAILS will conform to the requirements set by DANS EASY. Data users who want to use TRAILS data will need to comply with the terms and conditions set by DANS EASY, in addition to those set by TRAILS.

### **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*