



## Assessment Information

[CoreTrustSeal Requirements 2020–2023](#)

Repository:	Dutch Language Institute
Website:	<a href="https://www.ivdnt.org/">https://www.ivdnt.org/</a>
Certification period:	19 October 2023 - 18 October 2026
Requirements version:	CoreTrustSeal Requirements 2020-2022

This repository is owned by: **Dutch Language Institute**

## CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

### Background Information

#### Repository Type

Please provide context for your repository. You can select one or multiple options.

#### Compliance level:

Not Applicable - 0

#### Response:

- Institutional repository

#### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

Not Applicable - 0

#### Comments:

#### Reviewer 2:

#### Compliance level:

Not Applicable - 0

#### Comments:

### Description of Repository

Provide a short overview of the repository.

#### Compliance level:

Not Applicable - 0

#### Response:

The Instituut voor de Nederlandse Taal (Dutch Language Institute; IvdNT) is a binational institute, structurally funded by the Dutch and Flemish (Belgian) governments. It applies for the CTS as a CLARIN B Centre within the European CLARIN ERIC. The institute is actively improving its CLARIN services through the Dutch programme CLARIAH.

The institute collects and enriches Dutch linguistic material, and makes the material available as downloads and through software applications. The institute also accepts Dutch linguistic material collected and enriched by other institutes and universities, again made available through software applications.

#### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

Not Applicable - 0

## Dutch Language Institute

**Comments:**

**Reviewer 2:**

**Compliance level:**

Not Applicable - 0

**Comments:**

**Designated Community**

**Provide a clear definition of the Designated Community**

**Compliance level:**

Not Applicable - 0

**Response:**

The institute aims to serve anyone interested for whatever reason in the Dutch language. There is obviously a particular focus on researchers from the humanities and social sciences.

**Links:**

**Reviews**

**Reviewer 1:**

**Compliance level:**

Not Applicable - 0

**Comments:**

**Reviewer 2:**

**Compliance level:**

Not Applicable - 0

**Comments:**

**Level of Curation**

**Select all relevant types of curation.**

- Content distributed as deposited
- Basic curation – e.g., brief checking, addition of basic metadata or documentation
- Enhanced curation – e.g., conversion to new formats, enhancement of documentation
- Data-level curation – as above, but with additional editing of deposited data for accuracy

**Compliance level:**

Not Applicable - 0

**Response:**

- A. Content distributed as deposited
- B. Basic curation – e.g. brief checking; addition of basic metadata or documentation
- C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation

**Links:**

## Dutch Language Institute

### Reviews

#### Reviewer 1:

#### Compliance level:

Not Applicable - 0

#### Comments:

#### Reviewer 2:

#### Compliance level:

Not Applicable - 0

#### Comments:

### Level of Curation - explanation

Please add the description for your Level(s) of Curation.

#### Compliance level:

Not Applicable - 0

#### Response:

Our standard is to do A, B, and C. D will only be done when additional funding is available.

#### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

Not Applicable - 0

#### Comments:

#### Reviewer 2:

#### Compliance level:

Not Applicable - 0

#### Comments:

### Insource/Outsource Partners

If applicable, please list them.

#### Compliance level:

Not Applicable - 0

#### Response:

Most of the infrastructure is managed by the institute itself. We also create our own handles (prefix 10032). The servers are in a data centre of the University of Leiden. The IT department of the university (ISSC) takes care of the firewall, switching and cables. The institute has a long-standing relation with Centric IT Solutions and its partners, for consultancy and sometimes implementation. An important partner of Centric is BackupNed, that takes care of off-site storage of our back-up tapes in a bunker. The data and applications are made available through virtual machines that can easily be transferred

## Dutch Language Institute

to another CLARIN B Centre. References:

<http://issc.leidenuniv.nl/>

<http://www.centric.eu/EU/Default>

<http://www.backupned.com/en/home>

### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

Not Applicable - 0

#### Comments:

#### Reviewer 2:

#### Compliance level:

Not Applicable - 0

#### Comments:

### Significant Changes

#### Summary of Significant Changes Since Last Application if applicable.

#### Compliance level:

Not Applicable - 0

#### Response:

-

### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

Not Applicable - 0

#### Comments:

#### Reviewer 2:

#### Compliance level:

Not Applicable - 0

#### Comments:

### Other Relevant Information

You may provide other relevant information that is not covered by the requirements.

#### Compliance level:

## Dutch Language Institute

Not Applicable - 0

### Response:

During the previous review the INT was the national coordinator for the Dutch Language Union (DLU), The DLU represented both the Netherlands and Flanders. But the membership of the DLU has been discontinued. Instead Belgium has become a new CLARIN member (CLARIN BE) in September 2021, with the INT as leading NC partner.

Reference: <https://clarin-be.ivdnt.org/>

### Links:

### Reviews

#### Reviewer 1:

##### Compliance level:

Not Applicable - 0

##### Comments:

#### Reviewer 2:

##### Compliance level:

Not Applicable - 0

##### Comments:

## Organizational Infrastructure

### R1 Mission/Scope

**The repository has an explicit mission to provide access to and preserve data in its domain.**

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Response:

The Dutch Language Institute (INT) collects the latest Dutch words, updates important works such as the General Dutch Grammar (Algemene Nederlandse Spraakkunst or ANS) and makes technical terminology accessible by providing term lists. Apart from that, the institute takes a central position in the Dutch-speaking world (the Netherlands, Flanders, Suriname and the Netherlands Antilles) as a developer, keeper and distributor of scholarly and sustainable language resources. With these corpora, lexicons, dictionaries, grammars and other products it provides the necessary building blocks for the study of the Dutch language.

The mission and responsibilities of the repository are described in this document: <https://portal.clarin.inl.nl/information-about-deposition.html>

The repository is one of the main tasks of the institute as can be found in Article 2.1.c of the statutes of the INT:

<https://ivdnt.org/wp-content/uploads/2022/09/bijlage1-afschrift-statuten.pdf>

References:

<https://ivdnt.org/over-ons/over-het-instituut/#about-the-institute>

### Links:

### Reviews

#### Reviewer 1:

##### Compliance level:

The guideline has been fully implemented in the repository - 4

## Dutch Language Institute

### Comments:

### Reviewer 2:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### R2 Licenses

**The repository maintains all applicable licenses covering data access and use and monitors compliance.**

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Response:

Most of the data in the INT repository is property of other parties. If data is owned by third parties, like publishers, a provider agreement makes explicit all obligations and restrictions on use. The data producer, i.e. the provider will always remain the proprietor. INT does in fact get a copy of the data of which it must take good care, according to the terms of the license contract and the terms and conditions for use. INT also makes copies, for example for the benefit of back-up, and looks after them well. Provider agreements or contracts vary from product to product and are not made public. The provider agreement also specify under what conditions access to the data is provided. Some providers prohibit full downloads of the data, some prohibit commercial applications and some demand signed license agreements.

The end user license for the INT CLARIN services are explained in simple and clear wording in the "End User License Agreement" in our CLARIN portal. Before downloading is allowed, end users first are required to agree to the content of a license document they are presented with. We use mainly proprietary end user licenses which cover the agreements with the providers of the data.

Breaches of contract will be handled by the legal representative of the INT. This situation has not yet occurred.

### References:

[https://portal.clarin.ivdnt.org/doc/end\\_user\\_license\\_agreement\\_INT.pdf](https://portal.clarin.ivdnt.org/doc/end_user_license_agreement_INT.pdf)

### Links:

### Reviews

### Reviewer 1:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### Reviewer 2:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### R3 Continuity of access

**The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.**

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Response:

## Dutch Language Institute

The activities of the INT, including the data repository, are financed on a permanent basis by the Dutch and Flanders governments. Therefore, there is certainly no concern that the operations will have to desist in the short and medium term. Also for the long term there are no indications that the operations are in danger. Therefore we can offer depositors that their data will remain preserved and accessible indefinitely. This applies to all products in our catalogue. Nevertheless, the infrastructure has been set up according to mainstream concepts so that data and applications can be moved to another CLARIN B Centre with a minimum of effort.

In the Statute of the INT it is stated that in case of desistance, all assets of the INT, including the repository, will be transferred to the Dutch Language Union, which will decide on further management of the data. There are no explicit procedures in place for such a transfer, since it will depend on the receiving organisation.

References:

<https://ivdnt.org/wp-content/uploads/2022/09/bijlage1-afschrift-statuten.pdf>

It is difficult to find online information about the financial commitment of the Dutch and Flemish governments, but the accountant's statement happens to mention it at page 15 "Belangrijkste risico's": <https://ivdnt.org/wp-content/uploads/2021/06/Jaarrekening-2020-definitief.pdf>

**Links:**

### Reviews

**Reviewer 1:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Comments:**

**Reviewer 2:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Comments:**

### R4 Confidentiality/Ethics

**The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Response:**

When the INT ingests linguistic material, it uses a Standard Operating Procedure that also include checks on possible personal information that might be included in the data. Checks include inspecting the nature of the data by specialized staff and consulting the creators or compilers of the data. Data that contains personal information like speech or video of subjects is only digested if it is accompanied by the appropriate consent forms. The INT has an internal Data Protection Officer who monitors these processes. Data that contain personal information might be licensed with additional conditions depending on the type of consent that has been obtained.

The INT CLARIN B Centre is embedded in the INT. The INT is a recognized research institute and as such adheres to the Netherlands Code of Conduct for Academic Practice.

See [http://www.vsnul.nl/files/documenten/Domeinen/Onderzoek/The\\_Netherlands\\_Code%20of\\_Conduct\\_for\\_Academic\\_Practice\\_2004\\_\(version2014\).pdf](http://www.vsnul.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code%20of_Conduct_for_Academic_Practice_2004_(version2014).pdf)

**Links:**

### Reviews

**Reviewer 1:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4



## Dutch Language Institute

### Comments:

### Reviewer 2:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### R5 Organizational infrastructure

**The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.**

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Response:

The INT CLARIN B Centre is embedded in the INT. Its mission and funding have already been discussed. The INT and therefore the CLARIN B Centre have an abundance of trained staff: linguists (16), computational linguists (4), software developers (6), a systems architect and systems developers. The funding covers frequent participation of many staff members in relevant meetings and seminars. The quality and training of staff is a spearhead of the policy of the director of the INT. The INT has an Advisory Board, consisting of persons with extensive knowledge and experience of the work areas of the foundation. Members of the Advisory Council are nominated by the Supervisory Board and appointed by the Committee of Ministers of the Language Union. The Advisory Board gives asked and unsolicited advice to the Board and the Committee of Ministers of the Language Union on issues relating to the realization of the Foundation's objectives.

### References:

<https://ivdnt.org/over-ons/over-het-instituut/>

### Links:

### Reviews

### Reviewer 1:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### Reviewer 2:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### R6 Expert guidance

**The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either in-house, or external, including scientific guidance, if relevant).**

### Compliance level:

The repository is in the implementation phase - 3

### Response:

The INT-Taalmaterialen has a dedicated servicedesk which communicates directly with the designated community. In this way we regularly receive suggestions for improvement and requests from users. Beside that, the INT participates regularly in international projects in which we are required to work

## Dutch Language Institute

with new technologies which broadens our technological horizon. There is no dedicated advisory board for the repository.

The INT technical infrastructure is embedded within that of Leiden University. Therefore we have regularly contact with the IT department of the university (ISSC) pertaining to aspects like data security and access security.

The INT, including the repository service 'Taalmaterialen' is regularly evaluated by an independent international visitation commission. The last evaluation was in 2021. The result of the visitation was very positive. The report (in Dutch) can be found here:

<https://ivdnt.org/wp-content/uploads/2021/12/Visitatierapport-INT-NL.pdf>

### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

The repository is in the implementation phase - 3

#### Comments:

#### Reviewer 2:

#### Compliance level:

The repository is in the implementation phase - 3

#### Comments:

## Digital Object Management

### R7 Data integrity and authenticity

**The repository guarantees the integrity and authenticity of the data.**

#### Compliance level:

The repository is in the implementation phase - 3

#### Response:

To ensure the integrity of the data sets, for every deposited file a checksum (md5 type) is made which allows us to check the files for changes in later years. Checksums are calculated for every single file in a collection. Once deposited, files in data sets are never changed and only minor changes to the metadata are allowed. For example: correction of spelling, minor changes in documentation, additional documentation added, etc. Changes to the data themselves will be issued as a new version of the dataset and will obtain a new persistent identifier. In case of minor corrections to the data a subversion is issued. In case data is extended a new version is issued. These changes are only made in narrow collaboration with the provider of the dataset. The older versions are archived in the complete original form.

The correspondence with the producer and other relevant documents are archived with the dataset.

A register is maintained for each dataset which records details of the dataset, the ingestion process (dates, actors, etc) and subsequent changes (dates, actors, etc). These registers are, however, not public and can not be shared. Moreover, all correspondence with provider(s) is archived.

The repository maintains links to other relevant materials (e.g. article, thesis, documentation, data elsewhere) on the productpages of the datasets.

Example: <http://hdl.handle.net/10032/tm-a2-v4>

The repository doesn't compare the essential properties of different versions of the same file.

The unique identity of a depositor is ensured either through personal contact or by the required login using the CLARIN Service Provider Federation for identification.

#### References:

<https://portal.clarin.inl.nl/information-about-deposition.html>

### Links:

### Reviews

#### Reviewer 1:

## Dutch Language Institute

### Compliance level:

The repository is in the implementation phase - 3

### Comments:

### Reviewer 2:

### Compliance level:

The repository is in the implementation phase - 3

### Comments:

### R8 Appraisal

**The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.**

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Response:

The INT 'Taalmaterialen' repository only ingests materials that are relevant for Dutch language research and for the Dutch speakers community. The data is reviewed by INT staff, using the INT guidelines for deposition. Ref: <https://portal.clarin.inl.nl/information-about-deposition.html>

We provide CMDI metadata for all products in the repository. These metadata are published in the CLARIN VLO (<https://repository.ivdnt.org/oai/provider?verb=ListRecords&metadataPrefix=CMDI>).

We have a policy to provide data as much as possible in open standards. The data from other producers is accepted if it complies to these standards. If it does not comply we will assist the producer to comply to this policy. All datasets are checked by INT staff during ingestion.

When datasets or withdrawn from the repository, the PID that has been issued for that dataset will point to a tombstone page which explains why the data has been removed and, if possible, links to other locations where the data can be found.

### Links:

### Reviews

### Reviewer 1:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### Reviewer 2:

### Compliance level:

The guideline has been fully implemented in the repository - 4

### Comments:

### R9 Documented storage procedures

**The repository applies documented processes and procedures in managing archival storage of the data.**

### Compliance level:

The repository is in the implementation phase - 3

### Response:

## Dutch Language Institute

The management is taken care of by the IT staff of the INT and takes place according to extensive documentation and numerous Standard Operating Procedures and checklists, stored in an internal wiki.

The procedures include: Weekly full back-ups to tape. The full back-ups are followed by incremental back-ups on a daily basis. Quarterly full back-ups to tape, to be stored at another location (Backupned).

These tapes will eventually be reused, but one set of each year is retained for at least seven years. A restore can be carried out upon request. Essential information for disaster recovery is stored on paper and digitally in a vault at a great distance from the data centre.

Installation of security patches and updates on a monthly basis.

The INT internal Wiki contains a vast number of documents, concerning both ongoing projects and project results (applications, data). These concern CLARIN as an umbrella, software development and system management. The documents consist for the most part of documentation, standard operating procedures, checklists and best practices. Internal INT standards are applied also to external data and applications (metadata, security etc.). The information in the internal wiki is in Dutch and is confidential. The tables of contents of the most relevant sections can be sent separately to the reviewer if he or she wishes.

Older versions of datasets are archived and preserved in their complete form. Consistency over time is checked by registering checksums.

At the moment no risk management is used other than common sense, years of experience and adopting well-established practices.

### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

The repository is in the implementation phase - 3

#### Comments:

#### Reviewer 2:

#### Compliance level:

The repository is in the implementation phase - 3

#### Comments:

### R10 Preservation plan

**The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.**

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Response:

As much as possible open file formats are used. Small volume conversions due to obsolescence of file formats will be handled during standard maintenance cycles. Large conversion projects need to be handled as projects and might need additional funding. For textual resources, XML formats are used whenever possible, to make future interpretation of the files possible, even if the tool that was used to create them no longer exists.

The INT has the technical capacity to convert data when needed and follows the guidelines published by CCSDS, in the Reference Model for an Open Archival Information System (OAIS).

The OAIS presents a functional model consisting of six functional entities. A number of interactions are possible between those entities. We will present a description of these entities within the INT CLARIN Centre.

1. Ingest. This entity receives data from producers. Special tasks are: receiving data, performing quality assurance, checks on documentation, description and formats. Establish metadata and prepare for archiving and data management. Implications for INT CLARIN Centre: there is a Standard Operating Procedure for ingest of data (acquisition) which includes all the tasks mentioned.

2. Archival Storage. This entity is responsible for the systematic storage, maintenance and retrieval of the data. It further performs routine checks on media quality (refresh if necessary), errors and disaster recovery capabilities. Implications for INT CLARIN Centre: the INT CLARIN Centre distinguishes two separate functions. First, data management, which is responsible for storage of the data, error detection and retrieval. Second, system management, which is responsible for media quality and recoverability.

3. Data Management. This entity is responsible for content integrity of the data. It sees that data and descriptive information is connected and is responsible for version management. Implications for INT CLARIN Centre: these responsibilities are part of the previously mentioned data management

## Dutch Language Institute

function.

4. Administration. Oversees all archiving operations. Negotiates submission agreements with Producers. Establishes policies for maintenance, standards and hardware and software planning, customer support, etc. Implications for INT CLARIN Centre: the coordinator of INT CLARIN Centre, together with management of INT, is responsible for developing these policies.

5. Preservation planning. Evaluates the quality of the content and the quality of the service in context of the user community. Signals developments in technology and use patterns and provides policies to upgrade the archive service accordingly. Also provides migration planning. Implications for INT CLARIN Center: the INT does not have an explicit policy in this, but stays on top of technological developments by other means. The INT collaborates in projects with many national and international parties. The services of the INT CLARIN Centre are continuously updated according to the needs of those collaborations. Furthermore, the INT CLARIN Centre participates in spearheading projects like CLARIAH that intend to provide a digital infrastructure for research.

6. Access. This entity is responsible for the interaction with consumers. It provides information about the available products and is responsible for communication with consumers. Implications for INT CLARIN Centre: information about products is disseminated through a number of portals like OLAC, ELRA Universal Catalogue and CLARIN VLO. These services points interested parties to our products. Questions can be directed to the service desk.

### Links:

- [Preservation Plan](#)
- [guidelines published by CCSDS](#)

### Reviews

#### Reviewer 1:

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Comments:

#### Reviewer 2:

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Comments:

### R11 Data quality

**The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality- related evaluations.**

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Response:

Trained repository staff checks the completeness of documentation provided by the depositor. Metadata and a detailed description of the data is produced in concert with the depositor.

CMDI metadata is created and published using much used profiles. The CMDI profiles we use contain components describing the type of data, the annotations types available and possible applications of the data. The completeness of the metadata is checked by using the validation process offered by CLARIN.

The Repository Servicedesk collects feedback from users and provides updates of the data or documentation if necessary.

When possible the Repository Servicedesk keeps contact with the depositor in order to discuss questions from users and to report any issues concerning quality.

Product pages display all available information on the data, which may include documentation, quality issues and references to published papers about the data.

### Links:

### Reviews

## Dutch Language Institute

### Reviewer 1:

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Comments:

### Reviewer 2:

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Comments:

### R12 Workflows

**Archiving takes place according to defined workflows from ingest to dissemination.**

#### Compliance level:

The repository is in the implementation phase - 3

#### Response:

All main workflows in use are documented in a wiki. Examples are procedures for data ingest, data archiving, packaging and publication of data, creating PID's, version management and system management. System management is organized according to ITIL principles. The wiki is in Dutch and is not publicly available.

All processes pertaining to long-term preservation are currently under review and need to be described in more detail.

The procedures for ingest include checks on intellectual property and personal information. Also we discuss with the producers their future plans with the data like updates, and if there are any known problems or errors. Further, we like to know if the data is used or distributed by others.

#### Links:

### Reviews

### Reviewer 1:

#### Compliance level:

The repository is in the implementation phase - 3

#### Comments:

### Reviewer 2:

#### Compliance level:

The repository is in the implementation phase - 3

#### Comments:

### R13 Data discovery and identification

**The repository enables users to discover the data and refer to them in a persistent way through proper citation.**

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Response:

## Dutch Language Institute

The INT Taalmaterialen repository has a dedicated website (<https://taalmaterialen.ivdnt.org/>). All datasets have their own product page. The product page gives a brief description of the dataset and lists a number of properties of the dataset and recommends how to cite the dataset (including the PID created for the dataset). The website offers a search facility which allows the users to find datasets according to a number of parameters like data type, linguistic sub-field, commercial use, annotation types, etc.

All datasets also have a dedicated set of metadata (CMDI) which are available at the int OAI-service (<https://repository.ivdnt.org/oai/provider?verb=ListRecords&metadataPrefix=CMDI>).

These metadata are harvested by CLARIN and made available in the CLARIN VLO (<https://vlo.clarin.eu/>)

The INT provides PID's for every dataset in its repository. In case of updates (minor and major) new PID's are created. PID's of obsolete versions lead to tombstone pages which contain links to the most recent version. As PID we use handles that we manage ourselves with our own prefix (10032).

### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Comments:

#### Reviewer 2:

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Comments:

### R14 Data reuse

**The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.**

#### Compliance level:

The guideline has been fully implemented in the repository - 4

#### Response:

As much as possible open file formats are used. Conversions due to obsolescence of file formats will be handled. For textual resources, XML formats are used whenever possible, to make future interpretation of the files possible even if the tool that was used to create them no longer exists.

We publish a list of preferred data formats: <https://portal.clarin.inl.nl/information-about-deposition.html>. This list has been adopted from lists published by DANS and CLARIN and contains the formats and data types considered most relevant for the designated community. The list will be reviewed every few years.

The INT has the capability to convert when needed and follows the guidelines published by CCSDS ('<http://public.ccsds.org/publications/MagentaBooks.aspx>'), in the 'Reference Model for an Open Archival Information System (OAIS)' ('CCSDS 650.0-M-2').

We already mentioned the Component MetaData Infrastructure (CMDI).

### Links:

### Reviews

#### Reviewer 1:

#### Compliance level:

The guideline has been fully implemented in the repository - 4

## Dutch Language Institute

**Comments:**

**Reviewer 2:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Comments:**

### Technology

#### R15 Technical infrastructure

**The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Response:**

The core infrastructure has been maintained on the basis of proven technology. The process around managing the infrastructure are based on ITIL and in lesser cases on ASL. Soft- and hardware inventory are maintained in a CMDB database. Lifecycle management evaluate and test systems and software on cost-effectiveness, stability and flexibility. Resource management procedures do the same for available resources. Systems are configured in a redundant fashion to guard against hardware failure. Backups to tapes are used for long storage file-data backups. Although we do have proprietary systems at the moment (Vmware, Netbackup, Microsoft windows) we make it a policy to use Opensource as often as possible.

Developments are currently underway to improve the infrastructure: we are moving to a new datacenter for improved network facilities and docker is being tested as an alternative for ESX.

In respons to the Nagios/Icinga question: We tested Icinga together witch zabbix and libreNMS as alternatives for Nagios. We found no improvements in performance or new important functionality, other than cosmetic. In addition Nagios is still well maintained and has a large userbase and therefore, we did not switch. The usability of Nagios will be reevaluated via the lifecycle process in 2 years.

**Links:**

### Reviews

**Reviewer 1:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Comments:**

**Reviewer 2:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Comments:**

ICINGA is the replacement for Nagios for some time, therefore the reviewer sees signs of obsolescence here. For the next application please show that you can keep with software obsolescence.

#### R16 Security

**The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.**

**Compliance level:**



## Dutch Language Institute

The guideline has been fully implemented in the repository - 4

### Response:

Outside security relies on scans by the security department of Leiden University as hosting provider. If the scan is positive, Leiden University will send us a message and INT IT will take action to resolve it. They also maintain our outside firewall. The INT websites secured with https, TLS 1.2 and the latest webserver versions (apache and nginx).

Inside security relies on the virtual segmented network with firewalls in between. Since we rely on the services of Leiden University, we can not provide all of the requested information.

Virus threats are mitigated by TrendMicro desktop software and clamd mail-scan software.

Internal access requires Active Directory authentication, authorization and accounting. Access to remote desktop working is secured by a token based two-factor authentication system.

Security logs are stored off-server.

The security officer role is part of our system administrator function. No specific risk analysis tools are used.

The security measures have shown to be functioning well and have not changed in several years.

### Links:

### Reviews

#### Reviewer 1:

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Comments:

#### Reviewer 2:

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Comments:

### Applicant Feedback

#### R17 Applicant Feedback

**We welcome feedback on the CoreTrustSeal Requirements and the Certification procedure.**

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Response:

-

##### Links:

### Reviews

#### Reviewer 1:

##### Compliance level:

The guideline has been fully implemented in the repository - 4

##### Comments:

## Dutch Language Institute

**Reviewer 2:**

**Compliance level:**

The guideline has been fully implemented in the repository - 4

**Comments:**

Overall comment: The centre shows clear experience of running a certified repository. The main challenge is to keep the repository infrastructure and its description up to date and not to fall behind.