# Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository IDS Repository 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 22, 2015.

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:                        IDS Repository
Seal Acquiry Date:                 May. 22, 2015

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

Previously Acquired Seals:         Seal date:                April 9, 2013
                                   Guidelines version:       2010 | June 1, 2010

This repository is owned by:       **Institut für Deutsche Sprache (IDS Mannheim)**
                                   R 5, 6-13
                                   68161 Mannheim
                                   Germany

                                   T +49 621 / 1581 - 0
                                   F +49 621 / 1581 - 200
                                   E trabold@ids-mannheim.de
                                   W http://www.ids-mannheim.de/

# Assessment

## 0. Repository Context

### Applicant Entry

*Self-assessment statement:*

The CLARIN-D Resource Center of IDS-Mannheim (IDS-Repository: http://repos.ids-mannheim.de/) is part of CLARIN-D (Common Language Resources and Technology Infrastructure Deutschland) - a web and centers-based research infrastructure for the social sciences and humanities. The aim of CLARIN-D and its service centers is to provide linguistic data, tools and services in an integrated, interoperable and scalable infrastructure for the social sciences and humanities. The research infrastructure is rolled out in close collaboration with expert scholars in the humanities and social sciences, to ensure that it meets the needs of users in a systematic and easily accessible way. CLARIN-D is funded by the German Federal Ministry for Education and Research.

CLARIN-D is building on the achievements of the preparatory phase of the European CLARIN initiative as well as CLARIN-D's Germany-specific predecessor project D-SPIN. These previous projects have developed research standards to be met by the CLARIN services centers, technical standards and solutions for key functions, a set of requirements which participants have to provide, as well as plans for the sustainable provision of tools and data and their long-term archiving.

The IDS-Repository serves as an archive for corpora compiled for documenting the German language (spoken and written). This includes corpora compiled by the IDS as well as corpora compiled by external data providers. An overview of the currently available corpora is available at http://repos.ids-mannheim.de/corporaoverview.html. For more information on the overall mission of the IDS repository see DSA 4.

Within CLARIN-D this resource center is a certified center of type A. CLARIN distinguishes a number of different center types that have different impact for the language resources and tools infrastructure. Type B centers offer services that include the access to the resources stored by them and tools deployed at the center via specified and CLARIN compliant interfaces in a stable and persistent way.

Within CLARIN-D the following requirements hold for centers of type A (https://www.clarin.eu/node/3542) and are fulfilled by this resource center:

- Centers need to offer useful services to the CLARIN community and to agree with the basic CLARIN principles (own architecture choice, explicit statement about quality of service, usage of persistent identifiers, adherence to agreed formats, protocols and APIs).

- Centers need to adhere to the security guidelines, i.e. the servers need to have accepted certificates.

- Centers need to join the national identity federation where available and join the CLARIN service provider federation to support single identity and single sign-on operation based on SAML2.0 and trust declarations. In case all resources at a center are open, setting up a Service Provider is optional.

- Centers need to have a proper and clearly specified repository system and participate in a quality assessment procedure as proposed by the Data Seal of Approval or MOIMS-RAC approaches.

- Centers need to offer component based metadata (CMDI) that make use of elements from accepted registries such as CCR (http://www.clarin.eu/ccr) in accordance with the CLARIN agreements, i.e. metadata needs to be harvestable via OAI PMH.

- Centers need to associate PIDs records according to the CLARIN agreements with their objects and add them to the metadata record.

- Each center needs to make clear statements about their policy of offering data and services and their treatment of IPR (intellectual property rights) issues.

- Each center needs to make explicit statements to the CLARIN boards about its technological and funding support state and its perspectives in these respects.

- Centers need to employ activities to relate their role in CLARIN to the research community in order to guarantee a research based status of the infrastructure and allow researchers to embed their services in their daily research work.

- Centers that are offering infrastructure type of services need to specify their services for CLARIN and the terms of giving service.

- Centers are advised to participate in the Federated Content Search with their collections by providing an SRU/CQL Endpoint. This content search is especially suitable for textual transcriptions and resources.

A short overview of all requirements for centers of type B is also given in the form of a checklist (https://www.clarin.eu/content/checklist-clarin-b-centers).

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

The link https://www.clarin.eu/content/checklist-clarin-b-centers creates an error message. The requirements for type B are however included in the above linked documentation.

# 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:*

The IDS-Repository closely follows the recommendations for standards and tools for compiling language corpora [1, in German] issued by "Deutsche Forschungsgemeinschaft (DFG). The corpora compiled at the IDS are all available in digitized form, but differ widely in terms of available metadata, method of storage, and legal terms of use. Archiving these corpora typically involves further curation to complement and standardize metadata, unify dataformats, and ensure availability of legal information.
Following [1], the major requirements for accepting resources for long term archival are:

(a) Metadata: Every resource must be provided in a standardized format or an exhaustive documentation of the proprietary format. At least for the whole resource, a minimum set of Metadata in Dublin Core (DC:title, DC:description, DC:publisher and/or DC:creator, DC:legalStatus) must be provided. Moreover, comprehensive documentation describing - depending on the resource - provenance of data, procedure of curation, necessary tools, formats, and a bibliography of publications about the resource, must be provided.

If the resource consists of several parts, for example a collection of papers, provision of metadata for the individual parts in appropriate form is strongly encouraged. Ideally, these metadata are provided in CMDI, but other forms such as well documented comma separated tables, from which CMDI metadata can be generated are accepted.

(b) Quality Assurance: Only resources that comply with CLARIN guidelines or are created in peer-reviewed scientific projects (with respect to scientific and scholarly quality) are considered for deposit. The depositor is required to sign an agreement stating that these guidelines are met (see also DSA Guideline 5).

Data sharing and reuse is promoted by providing free access to the data (download, webservices) and metadata (via the OAI-PMH protocol). The CLARIN infrastructure contains software components such as the VLO (http://www.clarin.eu/vlo/) which enable users to browse and search through combined catalogs that contain metadata of all CLARIN repositories.

**References**

[1] Empfehlungen zu datentechnischen Standards und Tools bei der Erhebung von Sprachkorpora.
http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standa

## 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:*

The IDS Repository recommends to use formats listed in the CLARIN standard recommendations [1]. The encoding for textual sources (plain text, XML, etc.) should be Unicode. In addition, for spoken corpora, the following formats are currently accepted.

The FOLKER data format (Documentation in German[2], XML Schema[3])
The EXMARaLDA data format (Documentation [4], DTDs[5])

For other formats we offer advice for conversion. However, as a general principle we also archive digital data in their original format in order to minimize the risk of conversion loss.

**References**

[1] http://www.clarin.eu/recommendations

[2] http://agd.ids-mannheim.de/download/FOLKER-Datenmodell.pdf

[3] http://agd.ids-mannheim.de/download/Folker_Schema.xsd

[4] http://jtei.revues.org/142

[5] http://www.exmaralda.org/downloads.html#dtd

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

While the list of Clarin standard recommendations is thorough and well laid out, it was not updated since 2009. There have probably been some changes, like support for TEI?

## 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:*

Following general CLARIN standards, metadata for the IDS Repository must be provided in the CMDI format with unique references to the actual resources. Comprehensive documentation on how to create CMDI compliant metadata profiles and instances is available at http://www.clarin.eu/cmdi

The creation of metadata files (instances) can be performed with any standard XML Editor, e.g. the XML Editor ARBIL (https://tla.mpi.nl/tools/tla-tools/arbil/

) that comes with CMDI support. Additionally, a set of tools is provided that allow data producers to create new or adapt existing metadata to the CMDI standard. This includes customizable transformation scripts for converting existing metadata in a variety of formats (Dublin Core, generic XML, comma separated tables) to CMDI, and extracting metadata from text data.

The granularity of CMDI metadata and objects is chosen by the (meta)data producer. The IDS Repository itself is able to handle a high granularity of metadata and objects.

Metadata elements must be compliant to the standards set in CMDI. Since CMDI is a component based approach which allows (meta)data producers to create custom tailored metadata profiles there is no limit to the usage of established standards etc. In order to be visible and useable in the CLARIN infrastructure CMDI metadata added to the IDS Repository needs to contain a minimum set of attributes (linked to data categories stored in the Clarin-EU Concept Registry (CCR: http://www.clarin.eu/ccr) which is enforced by the quality checks as part of the automated ingest and delivery procedures of the IDS Repository.

This includes:

1. Validation against CMDI Schemas before the ingest.

2. Integrity check for all referenced data.

3. Generation of an actionable URL for all CMDI records and data, and registration of the URL in a handle system (http://hdl.handle.net/).

4. Validation based on the validation procedures of the underlying Fedora-Commons backend.

5. Validation of CMDI Records delivered by the OAI Provider, using the underlying validation of the Fedora-Commons PROAI provider.

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

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

*Minimum Required Statement of Compliance:*

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

## Applicant Entry

*Statement of Compliance:*

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

*Self-assessment statement:*

The mission of the IDS Repository is to serve as the repository of a CLARIN-D resource center. The mission of CLARIN-D is to provide "linguistic data, tools and services in an integrated, interoperable and scalable infrastructure for the social sciences and humanities" (http://de.clarin.eu/en/home-en.html). Therefore a repository in which data, tools and according metadata is archived on a long term basis has to be operated by such a resource center.

This mission is in line with the general mission of IDS-Mannheim (Satzung des Instituts für Deutsche Sprache, &2(1) [1]), which states: "The foundation pursues the purpose of scientifically researching and documenting the German language in its contemporary use and its more recent history. It cooperates with other national and international institutions with a similar goal, and also provides scientific services."

The IDS Repository is part of the CLARIN infrastructure and thus does not carry out promotional activities on its own, but is embedded into such activities on CLARIN-D and the European CLARIN level. These activities include but are not limited to:

• Providing comprehensive information on the CLARIN mission through websites (clarin.eu, de.clarin.eu).

- Operation and maintenance of the Virtual Language Observatory (VLO) which provides means to search for data/tools to the end user (based on the metadata provided by the resource centers/repositories that are part of CLARIN).

- Presenting data, tools and services provided by CLARIN on conferences.

- Organization of dissemination conferences that aim at getting in touch with the user communities of CLARIN.

- Organization of and contribution to training courses and (summer) schools.

**References**

[1] http://www.ids-mannheim.de/org/pdf/Satzung-IDS.pdf

## 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:*

The IDS Repository is not a legal entity on its own. It is run by IDS-Mannheim which is an institution governed by public law.


Depositors must sign an agreement stating that they respect IPR (Intellectual Property Rights) and privacy issues and that they own all necessary rights required to deposit the data. In particular, data must be anonymized when applicable. Users must confirm that they will use resources only in the intended way. The depositor can choose to make the data publicly available, restrict access to academics via AAI (Authentication and Authorization Infrastructure), or to restrict access to individual users.


Guidelines and model contracts are provided for both, depositors and users in Clarin Terms of Use [1]. The Terms of Service and Privacy Policy have been amended to clarify IDS Mannheim as the legal entity. Model contracts for Depositors are indeed tailored individually for each depositor.


Examples for the declaration of consent of interviewees in the FOLK Corpus are available:


Declaration of Consent FOLK audio recordings (in German) [2]

Declaration of Consent FOLK video recordings (in German) [3]


They restrict the terms of use of the recordings to very specific research contexts, and explicitly exclude dissemination to third parties.

**References**

[1] http://repos.ids-mannheim.de/tou.html (visited: April 30, 2015)

[2] http://repos.ids-mannheim.de/resources/EINVERSTAENDNISERKLAERUNG_FOLK.pdf

[3] http://repos.ids-mannheim.de/resources/EINVERSTAENDNISERKLAERUNG_FOLK_Zusatz_Video.pdf

## 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:*

The IDS-Repository runs on a 3-node visualization cluster hosted by the IDS Mannheim. The necessary storage is provided by a redundant storage system. The machines are housed in a modern data center that was completely overhauled in 2014. It provides redundant air conditioning and redundant uninterrupted power supplies. Maintenance of the systems is performed by a team of trained personnel.

Access to the virtual server is restricted by a firewall. The storage hardware and hardware for virtual machines is replaced at regular intervals to the latest state of art.

The IDS-Repository, that is data and operating system, is backed up Monday through Thursday with incremental backups. Full (4th) respectively differential (1st, 2nd, 3rd, 5th) backups are performed every fourth Friday. Backups have a retention period of three months and are stored on a dedicated backup server on disks.

In the future, the IDS anticipates to keep a mirror of the most valuable data with a 3rd-party (Mannheim University), but legal, technical, and financial issues still need to be settled.

The IDS-Repository virtual machine, the backup server and other critical infrastructure is monitored with Icinga (= network and service monitoring software).

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

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

*Minimum Required Statement of Compliance:*

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

## Applicant Entry

*Statement of Compliance:*

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

*Self-assessment statement:*

Measures are taken to enhance the chance of future interpretability of the data. The number of accepted file formats is detailed in DSA Guideline 2, to make future conversions to other formats more feasible. As much as possible open (non-proprietary) file formats are used. 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. Text is encoded in Unicode to ensure future interpretability.

When a particular file format is in danger of becoming obsolete, appropriate curation steps take place.

All resources in the IDS Repository (metadata and actual data) are equipped with a checksum, which is checked on a regular basis in coordination with the backup schedule described in DSA Guideline 6.

## 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:*

Technical Workflows: The IDS Repository uses Fedora Commons as an underlying repository system. The ingest workflows of the IDS Repository are built on top of the batch ingest utilities provided by Fedora Commons. As detailed in DSA 3, extensive technical validation and automated curation takes place for ingesting CMDI metadata and the underlying data.

Overall Workflows: The general goal of the IDS Repository is to sustainably archive linguistic resources (corpora and tools) compiled and developed at IDS-Mannheim together with their metadata. In addition, IDS-Mannheim aims at providing archival services to academic researchers and institutions according to its basic mission (Satzung des Instituts für Deutsche Sprache, &2(1)) [1]. Selection of resources, and decision about archival is governed by institutional best practices, balancing provenance, utility, and funding.

### References

[1] http://www.ids-mannheim.de/org/pdf/Satzung-IDS.pdf

## 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:*

The data provider retains all intellectual property rights to their data. The depositor must grant distribution rights to the IDS Repository and choose an access model (public, academic, individuals). Access models are provided by the repository and distribution rights are specified in the data provision contract. Enforcing licenses by data users in the case of misuse is conducted by the property rights owner.

Crisis management is based on the technical solutions described in DSA Guideline 6. In addition, the IDS Repository archives all metadata and data in such a way that they can be easily migrated to and mirrored at other CLARIN resource centers. All metadata and data have a persistent identifier (PID), and are stored as self contained XML files. Legal aspects of the process of relocating data to another institution is addressed by templates of license agreements provided in CLARIN.

## 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:*

Harvesting of metadata is possible via OAI-PMH. Local search facilities are provided on the basis of the search interface of Fedora Commons (http://repos.ids-mannheim.de/fedora/objects). In addition, all CMDI metadata are harvested by the OAI-PMH of the virtual language observatory (VLO: http://www.clarin.eu/vlo/), which provide a central starting point when searching for resources in the CLARIN infrastructure. For some resources "deep search" is supported by the means of the CLARIN Federated Content Search (http://www.clarin.eu/fcs) interface.

The IDS has acquired a Handle prefix and runs an own Handle server for persistent identifiers. The IDS anticipates to have their prefix mirrored by EPIC and is currently negotiating this issue with EPIC. The IDS Repository itself does not offer a persistent identifier service on its own but relies on the IDS Handle server. The usage of PIDs is mandatory for resources and their CMDI metadata in CLARIN thus all resources added to the repository can be referenced using PIDs.

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*

The negotiations refered to here and in DSA Guideline 6 were already ongoing at the time of your first DSA application - was there no progress?

## 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:*

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

*Self-assessment statement:*

The integrity of the data is ensured by the version control in the Fedora-Commons backend. Metadata is a data stream within the digital object, and as such is version controlled like object data. CLARIN propagates the idea of reproducible research. Thus updates/new versions of resources typically are equipped with a new PID. Only marginal changes to CMDI metadata are versioned without registering a new PID.

Part of the archiving workflow is the integrity check of the data and the metadata by the archive manager. This is done both manually and automatically. The metadata is parsed for syntactic correctness and manually evaluated for completeness and soundness. The object data is tested for syntactic correctness if possible. All datastreams and versions are equipped with a MD5 checksum, which is checked in coordination with the backups as described in DSA Guideline 6.

## 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:*

The repository in principle makes the original deposited objects available in an unmodified way, if the objects are in one of the accepted file types and encodings. In case of changes by the data producer, the repository creates a new digital object with a new PID. In the case that the repository has to change the data, e.g., because a file format becomes obsolete and superceded, the original data are kept.

The repository only accepts works from the original data producers, who are acknowledged as such by means of elements in Dublin Core, or equivalent elements with according CCR categories in CMDI. We use CMDI relations (depending on the profile) to link between objects within a collection, and providing links from objects to additional information. An example CMDI record for the "Mannheimer Korpus historischer Zeitungen und Zeitschriften" is available at: http://hdl.handle.net/10932/00-017B-E0F5-4DD7-4D01-F.

External deposits are only accepted after a due dilligence process involving a check of the identity of depositors and clarification of all legall issues along the lines described in DSA Guideline 5.

## 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:*

The repository complies with the OAIS reference model's tasks and functions[1]. Moreover, the repository uses the Fedora Commons software, which is compliant with the Reference Model for an Open Archival Information System (OAIS) due to its ability to ingest and disseminate Submission Information Packages (SIPS) and Dissemination Information Packages (DIPS) in standard container formats.

The data consumer has direct access to the archived objects via the web, provided that access requirements have been met.

A more detailed description of the IDS Repository Functional Architecture along the OAIS reference model and Ingest Pipelines is available in [2].

**References**

[1] Reference Model for an Open Archival Information System (OAIS), Recommended Practice, CCSDS 650.0-M-2 (Magenta Book) Issue 2, June 2012 http://public.ccsds.org/publications/archive/650x0m2.pdf

[2] Functional Architecture and Ingest Pipelines. http://repos.ids-mannheim.de/reposdescription.html

## 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:*

All CMDI metadata are provided without access restrictions according to CLARIN-D policies.

Part of the actual data is also provided without access restrictions, but a significant part is protected. For some data, a shibboleth account is necessary, for some data a personal account is necessary to get access to the data. For some data sets, explicit permission from the depositor is needed. For a large part of the data, the data consumer needs to agree with a code of conduct, which also contains licensing terms.

An example of a protected resource is DeReKo . Access to a large part of the actual data of DeReKo is only possible via COSMAS II (http://cosmas2.ids-mannheim.de/). In order to access DeReKo via COSMAS II an end user license agreement has to be signed (http://www.ids-mannheim.de/cosmas2/projekt/registrierung/). For some sub-corpora of DeReKo the access is further restricted to IDS-internal use only (see http://www.ids-mannheim.de/kl/projekte/korpora/archiv.html for a list).

Some smaller parts of DeReKo are also available for download. These are licensed under Creative Commons (CC-BY-SA), namely the Wikipedia corpora (wpd, wpd11, wdd11) and the corpus "Reden und Interviews" (rei) (see http://www.ids-mannheim.de/kl/projekte/korpora/verfuegbarkeit.html#Download). Further corpora that are available for download (mk1, mk2, bzk) are under a special license that allows for non-commericial scientific use only and prohibits their re-distribution.

## 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:*

Data depositors need to make sure that IPR and personality rights are respected in their deposited data. They specify an appropriate licence that data consumer need to accept. Data are protected by an AAI and only available when accepting the licence.

In addition, the IDS Repository requires data consumers to comply with the DFG code of conduct for good scientific practice [1].

**References**

[1] http://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/gwp/

## 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:*

The system does not allow ingest of data into the repository without the specification of access criteria and without providing an appropriate licence. These license conditions are displayed to users in the CMDI metadata and must be accepted before obtaining access to the data.

## Reviewer Entry

*Accept or send back to applicant for modification:*

Accept

*Comments:*