



Assessment Information

[CoreTrustSeal Requirements 2017–2019](#)

Repository:	HZSK Repository
Website:	https://corpora.uni-hamburg.de/hzsk/en/repository-search
Certification Date:	13 February 2019
This repository is owned by:	Universität Hamburg



HZSK Repository

Notes Before Completing the Application

We have read and understood the notes concerning our application submission.

True

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background & General Guidance

Glossary of Terms

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.

Repository Type. Select all relevant types from:

Domain or subject-based repository, Institutional repository, Research project repository

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

Comments

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

Brief Description of the Repository's Designated Community.

The HZSK Repository (<https://corpora.uni-hamburg.de/repository>, <http://hdl.handle.net/11022/HZSK-0000-0000-2C76-B-REPOSITORY>) at the Hamburg Centre for Language Corpora (HZSK, <https://corpora.uni-hamburg.de>) of the University of Hamburg in Germany is a domain based repository focusing on the domain of language corpora, i.e. collections of spoken and written texts provided for research and teaching purposes. As this research data is created within research projects, the HZSK Repository also serves as a research project repository for these projects. In particular, corpora containing spoken language data, multilingual data and data from lesser-resourced or endangered languages are integrated and disseminated to the academic community. The HZSK Repository is thus aimed at various research communities within language based humanities research, including linguistics, second language acquisition, multilingual studies and other disciplines analysing not the language itself but the content of recordings and texts, e.g. ethnology or cultural studies. The HZSK also accepts data from the University of Hamburg which is not entirely within the main focus, e.g. treebanks or written historical corpora, and thus also serves as an institutional repository.

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:
Accept

Level of Curation Performed. Select all relevant types from:

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

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Comments

Data can only be ingested into the HZSK Repository if it complies with the repository's requirements regarding the legal situation, the formats of data and metadata, and curation standards (cf.

<https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>). Ingests are only performed by the staff at the HZSK. Before data deposit, the completeness and consistency of data and metadata is automatically and manually checked, documentation is collected from the data owner which is later displayed on the resource's landing page, and a depositor's agreement is signed. Depending on the deposited data formats, the resource is then further prepared for ingest by the creation of various derived formats for web dissemination as well as standardized metadata formats (the CLARIN Component Metadata Infrastructure (CMDI), <https://www.clarin.eu/content/component-metadata>, and Dublin Core (DC)) to be provided via OAI-PMH. To ensure transparent curation at the HZSK, the data set is versioned with Git and curation tasks are managed with the project management system Redmine, including custom trackers for recurring curation and data integration tasks.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Outsource Partners. If applicable, please list them.

1) Gesellschaft für Wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)

The repository makes use of a common CLARIN PID service

(<https://www.clarin.eu/sites/default/files/pid-CLARIN-ShortGuide.pdf>) based on the Handle System (<http://www.handle.net/>) and in cooperation with the European Persistent Identifier Consortium (EPIC). The usage of PIDs is mandatory for resources in CLARIN, thus all resources added to the repository may be referenced using PIDs. CLARIN has a contractual relationship with GWDG concerning the provision of PID services via EPIC API v2. The services provided are stipulated as follows:

GWDG provides PID services for eight prefixes for up to 50.000 PIDs per year and two PID services for up to a million PIDs per year. These PID services based on the handle system include:

- A service for minting PIDs via EPIC API v2 with its current releases during the period of contract.
- All minted PIDs are additionally twofold replicated to mirror servers of the EPIC consortium.
- A permanent resolution service of minted PIDs is goal of the EPIC consortium. A resolution service for at least ten years is guaranteed by GWDG.
- The offer distinguishes base costs for these services, that include the minting of up to 50.000 PIDs per year and additional costs, that include the minting of up to 1.000.000 PIDs per year.
- This does not include software development for special requirements of the project.

This outsource partner offers relevant functionality for guideline 13: "The data repository enables the users to discover the data and refer to them in a persistent way through proper citation."

2) Forschungszentrum (FZ) Juelich GmbH

The CLARIN-D Ticketing System (OTRS), as part of the CLARIN-D Helpdesk (<http://support.clarin-d.de/>) which is operated by the HZSK, is hosted on a virtual machine by the Forschungszentrum Jülich. CLARIN-D has a contractual relationship with FZ Juelich concerning long-term storage and hosting of workspaces and virtual machines.

This outsource partner offers partially relevant functionality for guidelines: 10 "Workflows" (after ingestion) and "15 Technical Infrastructure" as well as 11 "Data quality" as the Helpdesk is used for all user feedback.

3) Regionales Rechenzentrum (RRZ) - Universität Hamburg

The HZSK Repository is hosted on a virtual machine provided by the RRZ Hamburg, which provides reliable backup strategies for all data according to its policies. For further information on the RRZ, please also refer to R9 and R16.

4) CLARIN-D

The repository is one of currently eight resource and service centres of CLARIN-D. As part of the CLARIN-D consortium, the repository has signed the "Kooperationsvereinbarung" - Cooperation Agreement - which states the rights and obligations of all CLARIN-D centres. A condensed version of this contract (in German only) is available at:

<https://www.clarin-d.net/de/ueber/zentren/zusammenarbeit>

CLARIN-D offers several services to its member institutions, among them the following:

- CLARIN-D HelpDesk (<https://support.clarin-d.de/mail/>): A central system for user support, which allows for the distribution of user questions and feedback to qualified personnel at the centres.
- CLARIN-D website (<https://clarin-d.de/en/>): A starting point for researchers to find information on CLARIN-D and to access CLARIN-D services.
- CLARIN-D wiki (<https://www.clarin-d.de/mwiki/index.php/Hauptseite>): A central platform for CLARIN-D-related staff.

- CLARIN central monitoring (<https://monitoring.clarin.eu/>): A monitoring service offered to all CLARIN-ERIC members and maintained by the resource centre Leipzig.

5) CLARIN-ERIC

CLARIN-D is a member of CLARIN'S European Research Infrastructure Consortium (ERIC). CLARIN-ERIC offers central services to its members and users, as stated here: <https://www.clarin.eu/value-proposition>

The services are available to all centres in the member countries of the CLARIN-ERIC (<https://www.clarin.eu/content/overview-clarin-centres>).

The most important services of the ERIC cover the search functionality for the German CLARIN centres:

- Virtual Language Observatory - VLO (<https://vlo.clarin.eu/>): CLARIN's central metadata-based search engine, which contains metadata of all German CLARIN-centres.
- Metadata harvester: The VLO is kept up to date using the metadata harvester run by the CLARIN- ERIC.
- Federated Content Search - FCS (<https://www.clarin.eu/contentsearch>): Optionally, centres can provide the actual data of their resources for this central content search.
- CMDI Component Registry (<https://catalog.clarin.eu/ds/ComponentRegistry>): CLARIN's registry for components and profiles according to ISO-24622-1.

In addition, CLARIN-ERIC offers several further services such as central registries, user statistics management and, as an official EUDAT community, access to advanced EUDAT services.

Reviewer Entry

Reviewer 1

Comments:
Accept

For your recertification, documentation concerning the type of organisations mentioned under R9 should also be provided. These are all part of the collaboration that the repository seeks certification for.

Reviewer 2

Comments:
Accept

Other Relevant Information.

The Hamburger Zentrum für Sprachkorpora (HZSK) is part of a network of research infrastructure centres located throughout Europe. The aim of these centres is to provide language data, tools and services in an integrated, interoperable and scalable infrastructure to researchers in the humanities and social sciences.

CLARIN (<https://www.clarin.eu/>) is an acronym for “Common Language Resources and Technology Infrastructure”. It is a research infrastructure that was initiated from the vision that all digital language resources and tools from all over Europe

and beyond are accessible through a single sign-on online environment for the support of researchers in the humanities and social sciences. The CLARIN infrastructure is fully operational in many countries, and a large number of participating centres are offering access services to data, tools and expertise.

In 2012, nine CLARIN member countries created CLARIN-ERIC (European Research Infrastructure Consortium), which is an international legal entity that governs and coordinates CLARIN activities. CLARIN-ERIC members are governments or intergovernmental organisations which pay an annual fee to support the development and maintenance of the CLARIN research infrastructure.

Germany is one of the founding members of CLARIN-ERIC and contributes to CLARIN-ERIC via CLARIN-D (<https://www.clarin-d.net/en/>). CLARIN-D is an acronym for “Common Language Resources and Technology Infrastructure Deutschland”.

The CLARIN-D Resource Centre HZSK is one of currently eight German CLARIN-D Resource and Service Centres which form a web and centre-based research infrastructure for the humanities and social sciences. The aim of CLARIN-D and its service centres is to provide language data, tools and services in an integrated, interoperable and scalable infrastructure for researchers in the humanities and social sciences and related disciplines. 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. The CLARIN-D Resource Centre HZSK is part of the CLARIN-D consortium 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 service centres, 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.

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

The following requirements hold for CLARIN centres of type B, and are fulfilled by this resource centre:

- Centres need to offer useful services to the CLARIN community.
- Each centre needs to refer to CLARIN in a visible way on its website.
- Each centre needs to make explicit statements about its funding support state and its perspectives in this respect.
- Each centre needs to make explicit statements about CLARIN compliant resources and services available at the centre.
- Each centre needs to make clear statements about their policy of offering data and services and their treatment of IPR issues.
- The centre has to implement the GÉANT Data Protection Code of Conduct (DP-CoC) for each of its federated Service Providers.

- Centres need to have a proper and clearly specified repository system and participate in a quality assessment procedure as proposed by the CoreTrustSeal.
- Centres need to adhere to the security guidelines, i.e. the servers need to have accepted certificates.
- Centres 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.
- Centres need to offer component based metadata (CMDI) that make use of elements from accepted registries such as the CCR in accordance with the CLARIN agreements, i.e. metadata needs to be harvestable via OAI-PMH.
- Centres need to associate (handle) PIDs with their metadata records. These PIDs should be suitable for both human and machine interpretation, taking into account the HTTP-accept header. Individual files (e.g. a text, zip or sound file) can be referred to with either the PID of the describing metadata record in combination with a part identifier or with another PID.
- Centres can choose to participate in the Federated Content Search with their collections by providing an SRU/CQL Endpoint.

A short overview of all requirements for centres of type B is also given in the form of a checklist (<https://office.clarin.eu/v/CE-2013-0095-B-centre-checklist-v6.pdf>).

The CLARIN mission is to create an infrastructure that makes language resources and technology available and readily usable to scholars of all disciplines, in particular the humanities and social sciences. In our age we are presented by many challenges as we deal with language in electronic formats, in spoken, written, and multimodal forms, as a carrier of information, as an object of study, and otherwise. The volume of texts and recorded spoken texts is enormous, and it is growing exponentially. The sheer size of this material makes the use of computer-aided methods indispensable for many scholars in the humanities and in neighbouring areas who are concerned with language material. CLARIN is committed to boosting humanities research in a multicultural and multilingual Europe, by facilitating access to language resources and technology for researchers and scholars across a wide spectrum of domains in the humanities and social sciences (Krauwer, 2008).

References:

Krauwer, S. (2008). CLARIN: Common language resources and technology infrastructure. Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)

Information on virtualization at the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/virtuelle-server/daten/virtuelle-server-ordnung-rrz220307.pdf>

Information on backup at the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup.html>

Backup-Policy of the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup/daten/tsm-nutzungsbedingungen.pdf>

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

ORGANIZATIONAL INFRASTRUCTURE

I. Mission/Scope

R1. The repository has an explicit mission to provide access to and preserve data in its domain.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

The general mission of the Hamburg Centre for Language Corpora providing the HZSK Repository as stated in its "Satzung" (articles of association) includes explicit references to both the tasks of archiving and making language resources available within existing and emerging digital infrastructures, and to promoting methodology related to these resources.

The first paragraph of the §2 of the Satzung describes the goals and tasks as a centre at the University of Hamburg ("The HZSK promotes and coordinates computer based research and teaching in linguistics and related disciplines at the University of Hamburg."), for which one of its main aims is "a. Ensuring sustainability, i.e. long-term usability and availability of empirical digital linguistic data, created and used at the University of Hamburg for research and teaching purposes".

The third paragraph of the §2 of the Satzung describes how the HZSK will participate in emerging digital infrastructures to allow new ways of accessing and working with digital resources, while the second paragraph of the §2 of the Satzung describes the role of the HZSK as a "competence centre" providing its expertise within the field of language corpora to scholars at the University of Hamburg and beyond. The HZSK thus presents its activities on a regular basis, organizes workshops and training courses to introduce people to the underlying methodology.

The mission of the HZSK Repository is also to serve as the repository of a CLARIN-D resource centre. 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" (<https://www.clarin-d.net/en/>). The mission of the HZSK CLARIN-D Repository is thus to ensure the availability and long-term preservation of resources in the field of Humanities and Social Sciences, to preserve the knowledge gained in research, to aid the transfer of knowledge into new contexts, and to integrate new methods and resources into university curricula. For an overview of the mission and goals of the CLARIN research infrastructure, see the publication by Erhard Hinrichs (national coordinator of CLARIN-D) and Stephen Krauwer (former executive director of CLARIN-ERIC), Hinrichs, E.; Krauwer, S. (2014). The CLARIN mission is also further described in R0.

This mission is supported by the infrastructure of the University of Hamburg and by the integration of the repository into the national and international CLARIN infrastructures. As part of the CLARIN-D infrastructure, it shares the CLARIN-D mission to provide linguistic data, tools and services in an integrated, interoperable and scalable infrastructure for the Humanities and Social Sciences (<https://www.clarin-d.net/en/about>), and is committed to play an active role in the development of CLARIN's repository infrastructure.

The HZSK is furthermore a centre with a large number of members (cf. <https://corpora.uni-hamburg.de/hzsk/en/members>) affiliated with or somehow related to the University of Hamburg and interested in supporting the work of the centre. §4 (Membership) of the Satzung declares "A person can become a member if she/he is actively engaged in research or teaching in scientific fields relevant to the centre and document her/his interest in participating in achieving the goals of the centre."

The HZSK is also collaborating closely with the Center for Sustainable Research Data Management (<https://www.fdm.uni-hamburg.de/en.html>), which is responsible for research data from all disciplines from the entire University of Hamburg and only provide basic/generic services. Researchers from the designated communities of the HZSK are thus encouraged to contact the HZSK to receive discipline or resource type specific support.

References:

HZSK-Satzung (in German): <https://corpora.uni-hamburg.de/pdf/Satzung.pdf>

Hinrichs, E.; Krauwer, S. (2014): The CLARIN Research Infrastructure: Resources and Tools for E-Humanities Scholars. In: N. Calzolari et al. (Eds.), Proceedings of the Ninth International Conference on Language Resources and Evaluation

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

II. Licenses

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

While all standardized metadata in CMDI and DC formats is made publicly available, access to most of the resources stored in the repository is restricted for reasons of data protection or IPR. The access restrictions for language resources hosted by the HZSK are based on the CLARIN License Categories: CLARIN PUB, CLARIN ACA or CLARIN RES (see <https://www.clarin.eu/content/license-categories> and <https://corpora.uni-hamburg.de/hzsk/en/corpus-enquiries-licenses>).

The authentication of users is carried out completely via Shibboleth (adding convenience for users and additional security), and users have to accept general Terms of Use after their first login. Most of the hosted resources can only be used for non-commercial research purposes, redistribution is not allowed, and special conditions apply for resources containing personal data.

Access to RES resources is only granted upon request on a case-to-case basis. Registered users need to explicitly

request access to individual RES resources by confirming and providing additional information about themselves (affiliation, academic status, full name, address etc.) and the intended use of the resource. All HZSK staff handling access requests have received dedicated training and use internal guidelines for the safe and confidential handling of personal data.

The specific conditions of use for restricted language resources can, however, not be the same for all RES resources in the repository, but are partly resource specific and therefore further specified in the contract/conditions of use accepted/signed by the data consumer for the individual resource. The End User Licence(s) used with data consumers thus depend on the rights holders' requirements, e.g. if desired, rights holders can require data consumers to sign contracts specifying the conditions of use for a particular resource in detail as defined by the rights holder.

For each hosted resource, an individual contract implementing individual access restrictions is set up with the rights holder(s) (in most cases the data producer(s)). All contracts used with the data consumers are based on one model contract with details on access that are adapted according to the requirements of the rights holder. The repository will not allow any deposit of data without a signed agreement specifying the handling of the data and access to it in detail.

While the HZSK can support the rights holder/data depositor in the curation process, e.g. with anonymization of files, the means of distribution of sensitive data is at the discretion of the rights holder, who is also responsible for the consent of the participants in the data and for compliance with ethical codes of conduct and national and international legal regulations. The conditions of use contain a part where the data consumer consents to not reveal the identity of any participant, nor publish data or part thereof in a manner that would make the reconstruction of a person's identity possible.

In case of non-compliance with the conditions or misuse, the user is denied further access to the repository. Further legal measures remain reserved to the data depositors.

Personal data that is retrieved from data consumers in the framework of the above-mentioned process or from users' home institutions via Shibboleth Login is processed according to the HZSK Privacy Policy:

<https://corpora.uni-hamburg.de/hzsk/en/privacy-policy>

References:

Example corpus access request form (for the Hamburg Corpus of Argentinian Spanish):

https://corpora.uni-hamburg.de/hzsk/en/hzsk_access/request/spoken-corpus:hacaspa (requires login)

Terms of Use

<https://corpora.uni-hamburg.de/hzsk/en/corpus-enquiries-licenses>

A copy of the model contract (in German) containing English comments:

https://corpora.uni-hamburg.de/pdf/DELA_HZSK_EN.pdf

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

III. Continuity of access

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The HZSK guarantees preservation of the deposited data for ten years, which is needed by research projects e.g. for compliance with "Recommendation 7: Safeguarding and Storing of Primary Data" of the DFG (German Research Council) guidelines "Safeguarding Good Scientific Practice". In case of cessation of funding within this period, strategies have been developed to cope with such a situation.

The medium-term plans rely on the granted funding for the CLARIN centre in the current project phase with the option of resorting to strategies set up for the long-term in the extremely unlikely event of a withdrawal of funding. In preparation for an event where immediate transferal of the repository and/or the resources would become necessary, daily incremental backups allow restoring the whole virtual machine on which the HZSK Repository runs. This is implemented according to the policies of the Computing Centre of the Hamburg University (Regionales Rechenzentrum, RRZ), which host the virtual server. The Fedora Commons system also stores digital objects in a transparent way, enabling transferal to other repositories. Since the data and metadata is stored in open standardized formats with a high degree of interoperability (CMDI, Dublin Core, EXMARaLDA, reIANNIS/PAULA), transferal of individual data sets to other technical infrastructures

is also feasible.

The long-term plans to ensure ongoing access and preservation includes two options for transferal of resources in the event of an unexpected cessation of funding. CLARIN centres commit to ensuring long-term availability, access and to preservation of datasets submitted to their repositories, as set out in their Mission statements. This is possible since CLARIN centres are set up as a distributed network, where each centre institution is a hub of the digital humanities and brings its own financial resources into CLARIN-D, which ensures continued availability. Accordingly, in the case of a withdrawal of funding, the repositories' content would be transferred to another CLARIN centre (cf.

<https://www.clarin-d.net/en/about/centres/mou-taking-other-centre-s-data>). The legal aspects of the process of relocating data to another institution is addressed by templates of license agreements provided in CLARIN. Transferal of resources to a CLARIN centre with a similar technical infrastructure and thematic focus would also allow similar methods of dissemination of the resources.

In case the rights holder prefers that the data remains at the University of Hamburg, the Center for Sustainable Research Data Management at the University of Hamburg (<https://www.fdm.uni-hamburg.de/>) will accept HZSK resources and host these in a more generic research data repository while continuously enabling reliable retrieval and citation of the resources: The Center for Sustainable Research Data Management (RDM Center) is a permanent central institution of the Universität Hamburg, which is assigned to the Vice-President for Research (VP3). The RDM Center bundles infrastructures, competencies and tasks in cooperation with the Regional Data Center (RRZ) in order to offer sustainable services. The RDM Center commits itself to preserve the digital assets stored in the HZSK repository in its own research data repository in case the HZSK goes out of service and other CLARIN centres are unable to take over the data or data owners wish for the data to remain at the Universität Hamburg. Persistent identifiers pointing to digital assets in the HZSK repository will still be resolvable.

In the unlikely event that the mission of the HZSK changes, i.e. by a modification of the Satzung (articles of association) that would e.g. restrict the thematic focus and the types of resources hosted further and make the transferal of no longer relevant resources necessary, both options can be applied as agreed upon with the rights holder/data depositor for individual resources.

References:

Center for Sustainable Research Data Management at the University of Hamburg:

<https://www.fdm.uni-hamburg.de/en.html>

Commitment of the Center for Sustainable Research Data Management (RDM Centre):

<https://corpora.uni-hamburg.de/pdf/FDM-Commitment.pdf>

DFG Guidelines "Safeguarding Good Scientific Practice":

http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf

Information on virtualization at the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/virtuelle-server/daten/virtuelle-server-ordnung-rrz220307.pdf>

Information on backup at the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup.html>

Backup-Policy of the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup/daten/tsm-nutzungsbedingungen.pdf>

CMDI component metadata: <http://www.clarin.eu/cmd/>

The EXMARaLDA formats: <http://www.exmaralda.org/>

ANNIS, PAULA and related formats: <http://corpus-tools.org/home/>

CLARIN-D Memorandum of Understanding "Taking over research data of other CLARIN-D centres":

<https://www.clarin-d.net/en/about/centres/mou-taking-other-centre-s-data>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

IV. Confidentiality/Ethics

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The HZSK Repository contains mainly spoken language corpora. For research projects starting to compile corpora, the HZSK provides assistance in all areas related to not only data formats and workflows, but also to legal and ethical issues and compliance to regulations and requirements given in CLARIN.

For projects that are finished with their data collection, the HZSK supervises a curation process carried out by the project to ensure that data and metadata comply with the repository requirements. Before such a curation process starts, the legal status of the resource needs to be clarified and information on how the data was originally created is demanded. Furthermore, compliance with internal and CLARIN guidelines concerning scientific and scholarly quality and a contract about the means and procedure of accessing the data is necessary.

For projects and persons interested in data deposit, we provide information and links to various guidelines on our website: <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>.

As a part of the individual contract implementing individual access restrictions set up with the rights holder/data depositor (cf. R2), the HZSK Repository requires a statement ensuring that no (privacy or IPR related) rights of any third party will be infringed by using the deposited resource according to the conditions agreed upon. The rights holder is solely responsible for the consent of the participants in the data and for compliance with ethical codes of conduct and national and international legal regulations. The repository will not allow any deposit of data without such a signed agreement specifying the handling of the data and access to it in detail.

While the HZSK can support projects creating spoken language corpora, e.g. through providing information on common disclosure risks for this resource type, such as rich metadata for speakers of small speaker communities or the content of interview data, or through assistance with anonymization of files, the means of distribution of disclosive or potentially disclosive data remains at the discretion of the rights holder. Access to such resources is usually only granted upon personal request on a case-to-case basis according to the conditions agreed upon with the rights holder. This means, after authenticating via Shibboleth and accepting the general Terms of Use to register as users of the HZSK Repository, users need to request access to an individual resource while providing additional information on themselves (affiliation, academic status, full name, address etc.) and stating the intended use of the resource. Access to corpora in the repository is only possible if access has been granted. All data is managed, stored and distributed with great care. All HZSK staff handling potentially disclosive data and related access requests have received training and use internal guidelines.

For end users, the terms of use also contain a section in which the data consumer consents to not reveal the identity of any participant, nor publish data or part thereof in a manner that would make the reconstruction of a person's identity

possible. The HZSK also provides individual guidance in any case where a user is in doubt regarding disclosive or potentially disclosive data. In case of non-compliance with the conditions or misuse, the user is denied further access to the repository. Further legal measures remain reserved to the data depositors.

References:

A copy of the model contract (in German) containing English comments:

https://corpora.uni-hamburg.de/pdf/DELA_HZSK_EN.pdf

General Terms of Use: <https://corpora.uni-hamburg.de/hzsk/en/corpus-enquiries-licenses>

CLARIN-D Legal Helpdesk (reference to internal Legal Information Platform):

<https://www.clarin-d.net/en/training-and-helpdesk/legal-helpdesk>

Guidelines for data deposit (in German):

https://corpora.uni-hamburg.de/pdf/Leitfaden_Aufbereitungsaufwand_und_Nachnutzbarkeit_von_Korpora.pdf

DFG Guidelines on technical aspects (in German):

http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standards_sprachkorpora.pdf

DFG Guidelines on legal issues (in German):

http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standards_recht.pdf

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

V. Organizational infrastructure

R5. 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:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The Hamburg Centre for Language Corpora (HZSK) is a research data centre at the University of Hamburg focussing on language corpora that was founded in 2011. Since the HZSK is also a type B CLARIN centre, the repository staff members have access to training on data management, metadata, long-term preservation and professional development (offered by CLARIN-D and CLARIN-ERIC). This includes regular developer meetings, mobility grants for sharing of expertise, conferences, meetings with their respective scientific communities (called discipline-specific working groups) as well as a centralized knowledge base (user guide, wiki, bugtracker and mailing lists). CLARIN has a wide field of expertise in its collaborative network of centres, which come from within their respective fields of digital humanities.

Since the HZSK Repository is also part of CLARIN-D, a research infrastructure to support the sharing, use and sustainability of language data and tools for research in the humanities and social sciences, it benefits from the information and materials provided by CLARIN-D on a wide range of topics, including teaching material, help on data management plans and other discipline-specific support.

By being part of the CLARIN-D consortium, the repository gains access to funding for running and further developing a sustainable repository and resource centre to support these goals. Besides staff resources this includes a budget for attending national and international meetings such as conferences, workshops or internal developer meetings and meetings with the discipline-specific working groups.

Currently CLARIN-D is funded by the Bundesministerium für Bildung und Forschung (BMBF). The current project phase has a runtime of four years and is funded until September 30th 2020. As an alternative to project based funding, CLARIN-D currently pursues a permanent continuation of funding.

Apart from the integration into CLARIN-D, the HZSK is also integrated into the local research data infrastructure of the University of Hamburg. Workflows, data standards and technical solutions are being adapted and developed in close collaboration with the Centre for Sustainable Research Data Management (<https://www.fdm.uni-hamburg.de/en.html>) to ensure sustainability of resources and services. While the staff of the HZSK has been funded by CLARIN-D and similar infrastructure projects since 2011 with a maximum of 4.5 FTE, and 3 FTE are currently funded until 2020, the services of

the HZSK are thus also being integrated into the emerging permanent research data infrastructure of the more recently founded Centre for Sustainable Research Data Management of the University of Hamburg.

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

VI. Expert guidance

R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

As a CLARIN B Centre, the HZSK undergoes assessment by the CLARIN Centre Assessment Committee (<https://www.clarin.eu/governance/centre-assessment-committee>) on a regular basis, usually corresponding to the DSA/CTS assessment.

The repository, through its membership in CLARIN-D, is supported by external advisory committees. The International Advisory Board (IAB), CLARIN-D's scientific advisory board, is a group of CLARIN-D external experts who are consulted on new developments and discuss strategic and content related developments, also with a bird's-eye view of other developments in the communities. With experienced experts from various backgrounds, a high-profile international committee was formed for this purpose. Members of the IAB are currently: Reinhard Althenhöner, Christiane Fellbaum,

Björn Granström, John Nerbonne, Heike Renner-Westermann and Achim Streit. In addition to the regular reporting, CLARIN-D also invites the IAB to review its work in a dedicated meeting each year.

The joint Technical Advisory Board (TAB) of CLARIN-D and DARIAH-DE, is a committee that supports collaboration on the fundamental technical level between the two large research infrastructures for the humanities and social sciences. The issues of the Collaboration are: questions of technical protocols, infrastructural requirements on the level of archiving, interconnection, search, etc. Based on requirements, small working groups (for example on persistent identifiers, authorization and identification) are being formed in areas with an overlap of requirements. This avoids duplication of developments and allows an increased efficiency in implementation, but also interoperability where overlaps exist. This includes for example an option to grant access to one infrastructure for users of the other. Members of the Technical Advisory Board are currently: Jan Haji (Prague Institute of Formal and Applied Linguistics, CLARIN Centre), Margareta Hellström (ICOS Carbon Portal staff member), Peter Leinen (German national library, head of information technology), Karlheinz Mörth (Austrian Academy of Sciences), Wolfgang Nagel (Technical University of Dresden, Head of the centre of information services and high performance computing), and Sabine Roller (University of Siegen, head of the centre of information and media technologies).

CLARIN is committed to boosting humanities research in a multicultural and multilingual Europe, by facilitating access to language resources and technology for researchers and scholars across a wide spectrum of domains in the humanities and social sciences (HSS). To reach this goal and to contribute to overcoming the traditional gap between the Humanities and the Language Technology communities, we established an active interaction with the research communities in HSS in so-called discipline-specific working groups.

These groups act as a link between the CLARIN-D resource centres and the research communities which represent the users of the CLARIN-D infrastructure. Currently eight working groups act as consultants for the needs of the humanities, social sciences and particular disciplines. All together they consist of more than 200 academic professionals. Their main role is to advise CLARIN-D during the development and implementation of the infrastructure so that these efforts can best meet the needs of all research communities involved. The working group chairs further coordinate dissemination and best practice using CLARIN-D services in their member communities.

CLARIN-D organizes joint activities of the working groups. This includes the organization of working group meetings, organization of specialized and interdisciplinary workshops and the creation of joint reports. Further, communications between CLARIN-D centres and the working groups as well as groups among themselves are coordinated. Virtual meetings are held on a bi-monthly basis. Activities of the WG are described on the CLARIN-D Website (<https://www.clarin-d.net/en/disciplines/>). For communication, mailing lists and wiki contents are maintained. The HZSK is associated with the Working Group "Linguistic Fieldwork, Ethnology, Language Typology". Direct communication with the working group mainly takes place in meetings with the members, e.g. in jointly organized workshops.

Since the HZSK is a centre at the University of Hamburg with general member meetings and a board of directors, relevant input from the designated communities regarding the HZSK Repository can be gathered through these channels. There is a member meeting twice a year (or more often if required by five members or two directors) and the board of directors

meets in preparation of the member meetings and whenever required to discuss strategies and activities of the HZSK.

Furthermore, the HZSK interacts with several external research projects while providing support in the earlier stages of the research data lifecycle, and regularly presents the HZSK Repository and its resources at conferences and in workshops.

The HZSK can also be directly contacted via the CLARIN-D Help Desk

(<https://support.clarin-d.de/mail/?queue=HZSK&lang=en>). These channels provide valuable input from the designated communities.

References:

Checklist for CLARIN B Centres:

<https://www.clarin.eu/content/checklist-clarin-b-centres>

CLARIN-D Working Group 3 (WG3) "Linguistic Fieldwork, Ethnology, Language Typology":

<https://www.clarin-d.net/en/disciplines/linguistic-fieldwork-anthropology-language-typology>

HZSK members and board of directors:

<https://corpora.uni-hamburg.de/hzsk/en/members>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

DIGITAL OBJECT MANAGEMENT

VII. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The HZSK Repository does not allow anonymous depositing of resources. Before a corpus is integrated, the HZSK will be in contact with the rights holder to assess the data using the guidelines for data deposit and plan the curation process. The level of curation necessary is decided on according to the results obtained through validation of the resource, including manual and automatic checks for completeness and formal correctness of data and metadata, which is necessary for the deposit. Before the first ingest, version control is implemented with Git, while Redmine is used to plan, monitor and document any common and resource specific changes to the deposited data and metadata necessary for the ingestion. After ingest, Fedora Commons' ability to automatically generate checksums for ingested resources and keep audit trails is used.

The HZSK Repository only allows manual versioning of resources on corpus level. New corpus versions are created and ingested manually by the HZSK staff when required due to major data changes, such as highly relevant corrections or completions, further annotation layers etc. and the changes for a new version are also controlled by Git versioning and planned, monitored and documented with Redmine. The standardized CMDI metadata provided for all resources contains version information on the resource including related changes. Superseded versions of resources remain available on request using their PIDs with information about and a link to the newer version added to the landing page.

Metadata for resource description in the format of the EXMARaLDA Corpus Manager tool (Coma), used to manage metadata on corpora, communications/sessions, speakers and physical files of the corpus, is included for all spoken corpora, which are based on a Coma file. Via the CMDI metadata provided for all hosted resources, links between metadata records within a resource or from metadata to digital objects are provided.

References:

Guidelines for data deposit (in German):

https://corpora.uni-hamburg.de/pdf/Leitfaden_Aufbereitungsaufwand_und_Nachnutzbarkeit_von_Korpora.pdf

Slides from the talk "Corpus Workflows: Formalization and Operationalization of Workflows for the Creation and Publication of Language Corpora" describing workflows used at the HZSK and in the associated project INEL:

https://inel.corpora.uni-hamburg.de/wp-content/uploads/jettka_hedeland-2018-HZSK_INEL_Workflows.pdf

Corpus formats and standards: <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>

CMDI: <https://www.clarin.eu/cmdl>

Git: <https://git-scm.com/>

Redmine: <https://www.redmine.org/>

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

VIII. Appraisal

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

The HZSK Repository provides information on its policies for data selection, requirements for hosting and preferred formats for data and metadata (including schemata) on its website. It also provides a guideline for the cost-benefit analysis for the curation of legacy data to make its decisions whether to accept data transparent.

The HZSK Repository accepts spoken language corpora in the EXMARaLDA formats, or in formats that can be automatically converted into the EXMARaLDA formats. Depending on data complexity etc., other standardized best practice formats for the transcription data can be accepted. Written corpora to be disseminated as zip archives are

accepted in standardized and well documented formats, preferably convertible to the relANNIS format for dissemination via the ANNIS platform.

Metadata in the formats of the CMDI metadata profiles used for the data in the HZSK repository is required for ingest. For depositors, standardized CMDI (and Dublin Core) metadata can be created on the basis of the EXMARaLDA's Corpus Manager tool (Coma) format by automatic conversions at the HZSK. It is also possible to provide the metadata directly in the CMDI format using the CMDI metadata profiles developed for the HZSK repository, or, depending on the specific project context and agreements between the depositor and the HZSK, in other automatically processable formats from which the required CMDI metadata can be generated.

The HZSK does not allow corpora to be deposited without successful validation. The data depositor is informed about existing problems that need to be solved before deposit. Before ingest is possible, further necessary curation and conversion takes place in cooperation with the data depositor. The HZSK ensures compliance with the format requirements by using a corpus validation module based on the EXMARaLDA system. For other best practice formats, corresponding validation of the data is required. As part of the deposit workflow, we try to obtain as much relevant metadata and detailed information about data provenance as possible from the data depositor. Sufficient metadata is treated as a prerequisite for corpus curation, as described in the guideline. While we advise depositors to include as much metadata as possible, especially for legacy data, which might still be of very high relevance, it is not always possible to provide metadata according to a specified schema with required fields, even if these only cover general information. Such resources can be accepted as long as the metadata is sufficient for the interpretation and re-use of the data. We also ask for documentation used in the creation process, such as transcription and annotation guidelines, and relevant reference publications in which the resource and analyses based upon it are described.

While the goal of the curation process for resources is always the same (a consistent high quality corpus of EXMARaLDA or other standardized best practice formats with CMDI metadata for ingest into the repository), the process is partly highly individual depending on the source data. For the common parts however, we use Redmine to enforce recurring workflows for deposit, curation and ingest phases.

References:

Corpus formats and standards: <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>

Guideline for the cost-benefit analysis for corpus curation (in German):

https://corpora.uni-hamburg.de/pdf/Leitfaden_Aufbereitungsaufwand_und_Nachnutzbarkeit_von_Korpora.pdf

CMDI component metadata: <http://www.clarin.eu/cmdl>

EXMARaLDA Corpus Manager: <http://exmaralda.org/en/corpus-manager-en/>

The EXMARaLDA formats: <http://www.exmaralda.org/>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

IX. Documented storage procedures

R9. The repository applies documented processes and procedures in managing archival storage of the data.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase

Accept

Reviewer 2

Comments:

3 – The repository is in the implementation phase

Response:

Data that is accepted - after the first data audit and assessment of available information on data provenance and the legal situation according to the guidelines for data deposit - is transferred to servers of the University of Hamburg managed by the HZSK and put under version control using Git. The versioning makes any changes to the data after deposit transparent and reversible. Simpler cases of editing, completion or conversion can be done by trained student staff of the HZSK under guidance of regular staff. The final preparation for ingest and the ingestion process is however only conducted by dedicated regular technical staff. All data in the repository is also only managed, stored and distributed with great care by regular staff of the HZSK. Only technical staff explicitly responsible for repository management have corresponding access rights.

Workflows for the reoccurring tasks of the functional units of the repository (Ingest, Archival Storage, Data Management, Administration, Access) are formalized and applied as custom workflows in the Project Management System (Redmine). The operations include and combine (semi-)automatic and manual tasks. The explicit formulation of the workflows allows for a certain degree of independence from (the implicit competence of) individual staff members. Apart from the

formalisation as custom trackers, the workflows are also documented in detail using the Redmine wiki, which is not public but only available to HZSK staff.

In addition, the formal handling of tasks in the Project Management System (Redmine) is also used for repository maintenance which makes it possible to make temporal and administrative plans for checking for new software versions, compatibility issues, and other potential issues for the repository in regular intervals.

The repository data is ingested into and managed by the Fedora Commons system, which ensures persistent data storage (application-independent operating system-level access to files) and keeps track of the data integrity by means of fixity checks (checksum calculation). Data integrity is thus partly supervised by functionality provided by the repository software. Fedora versions datastreams, creates checksums, and has an audit trail for tracking changes to the content.

In addition to the functionality provided by Fedora, unintentional and malicious changes in the file system are addressed by regular backups made by the computing centre and additional auditing software. The repository runs on a virtual machine hosted by the Computing Centre of the Hamburg University (Regionales Rechenzentrum, RRZ), which means the HZSK relies on their routines for backup and risk management in this aspect. Daily incremental backups allow for a restoration of the whole server within a duration of 24 hours. Bitstream-conservation is provided by the Computing Centre of the Hamburg University, for a timespan not less than 10 years.

The HZSK provides access to curated spoken corpora from research projects and similar contexts as a centre of the University of Hamburg. While the HZSK fulfills a highly specific purpose, there are other institutions at the University of Hamburg with whom the HZSK collaborates closely to ensure synergy and avoid parallel structures. The Center for Sustainable Research Management at the University of Hamburg (<https://www.fdm.uni-hamburg.de/en.html>) is responsible for preserving the original project data (from projects of the University of Hamburg) for ten years to ensure good scientific practice. Furthermore, the FDM Centre and the HZSK collaborate with the Hamburg University Archive (<https://www.archiv.uni-hamburg.de/>, only available in German) to allow for archival appraisal of (analogue) legacy research data, which is not archived by the HZSK. As these two institutions have only recently been established, workflows and responsibilities are currently being formalized and documented.

References:

Guidelines for data deposit (in German):

https://corpora.uni-hamburg.de/pdf/Leitfaden_Aufbereitungsaufwand_und_Nachnutzbarkeit_von_Korpora.pdf

Redmine Workflows used in managing data:

<https://gitlab.rrz.uni-hamburg.de/hzsk-open-access/redmine-custom-workflows>

Information on virtualization at the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/virtuelle-server/daten/virtuelle-server-ordnung-rrz220307.pdf>

Information on backup at the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup.html>

Backup-Policy of the Computing Centre (RRZ) of the University of Hamburg:

<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup/daten/tsm-nutzungsbedingungen.pdf>

Reviewer Entry

Reviewer 1

Comments:

Accept

For your recertification, documentation about the workflows and responsibilities of the Center for Sustainable Research Management, the FDM Centre and the Hamburg University Archive should also be referred to at the beginning of the application, to support the information about outsourcing. These are all part of the collaboration that the repository seeks certification for.

Reviewer 2

Comments:

Accept

X. Preservation plan

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase

Reviewer 2

Comments:

3 – The repository is in the implementation phase

Response:

The HZSK repository follows the OAIS reference model in its conceptual and technical approach to long-term preservation and storage of the digital objects that are received, maintained, and delivered.

Guidance for Data Producers:

Prior to negotiating and initiating the submission of Submission Information Packages (SIP) which at least have to contain data objects, standardized metadata, and additional descriptive and procedural (presentation) information, depositors (typically linguists and/or linguistic research projects) are asked for self-evaluation and measurement of the quality and value of their data regarding the potential for re-usability and long-term preservation by consulting best practice guidelines for data deposit which are currently being transformed into a multilingual interactive web application). The guidelines include information about preferred data formats, necessary documentation, legal rights of subjects in the data, etc. In addition, the repository management at an early stage shares drafts of contracts for data usage and data transfer with the depositor to make sure that legal rights are understood and clarified. When depositing data with the HZSK Repository, the data depositor retains all rights to the deposited data, and the HZSK is only granted distribution rights and the right (and explicit obligation) to adapt the data to evolving technical standards. The depositor also decides on a license and the means of distribution for the resource according to the licenses in use at the HZSK, as specified in the individual contract. If no major conceptual concerns arise for the depositor and repository management, additional automatic and semi-automatic quality checks and fixes ("HZSK-Corpus-Services") are applied to the SIPs. This can result in revision of the data by the depositor, but ensures that a high degree of data consistency can be reached.

In order to be able to guarantee future adaption of deposited data through conversion and/or migration to future applications, we harmonize all data to make sure we only have current best practice formats and well-documented semantics of e.g. transcription data. This way, migration or conversion becomes feasible since one method can be used on all deposited data of a certain type. On our website with information on hosting and data deposit, we give data producers information about this as the reason for preferring certain formats and requiring a curation of data if it is to be reliably preserved. In case the curation cannot be conducted and data does not meet our requirements, preservation cannot be guaranteed. The HZSK is involved in defining standards and best practices for the specific area of audiovisual language corpora through as documented e.g. as DFG standard guidelines. These guidelines also provide the foundation for data preservation through harmonization and possibly future migration or conversion to be able to preserve deposited data regardless of changing standards over time.

Guidance for functional entities:

Workflows for the reoccurring tasks of the functional units of the repository (Ingest, Archival Storage, Data Management, Administration, Access) are formalized and applied as custom workflows in a Project Management System (Redmine) [2]. The operations include and combine (semi-)automatic and manual tasks. The explicit formulation of the workflows leads to a certain degree of independence from (the competence of) individual members of the staff. In addition, the formal handling of tasks is also used for repository maintenance which makes it possible to make temporal and administrative plans for checking for new software versions, compatibility issues, and other potential issues for the repository in regular intervals.

Data integrity is supervised by functionality provided by the repository software. Fedora versions datastreams, creates checksums, and has an audit trail for tracking changes to the content. In addition, unintentional and malicious changes in the file system are addressed by regular backups made by the computing centre and additional auditing software. The

online availability of the repository services is monitored locally by server-side scripts and centrally, e.g. by the CLARIN central monitoring.

Guidance for Consumers:

The Designated Community of the repository is typically seen as researchers or students of linguistic studies which have a certain degree of understanding for linguistic data formats and conventions. However, the repository presents as much information in its Dissemination Information Packages as possible, e.g. in the form of human-readable metadata and additional descriptions rendered as HTML. Much of the content in the repository can be viewed directly in its online interface, e.g. PDF views, HTML visualization for linguistic transcriptions. Context-sensitive help pages are reachable from each web page to guide users in the repository usage, and for further information users are directed to the CLARIN Help Desk to be found in a prominent place on each web page.

The preservation plan also ensures that in case of emergency deposited data can be transferred to another CLARIN centre with a similar profile or, if the rights holder prefers that the data remains at the University of Hamburg, to the recently founded Center for Sustainable Research Data Management, which will accept HZSK resources and host these in a more generic research data repository while continuously enabling reliable retrieval and citation of the resources (cf. R3). To ensure that a transfer of resources is feasible, all main corpus files in the HZSK repository are stored in open standardized XML-based formats, allowing for easy conversion into other formats. As outlined in R3, there are also standards for access models (PUB, ACA, RES) within CLARIN that apply across repositories.

As outlined in R9, there is a close collaboration with further institutions at the University of Hamburg responsible for archiving original project data to ensure good scientific practice and archival appraisal of analogue legacy data.

References:

Best practice guidelines for data deposit:

https://corpora.uni-hamburg.de/pdf/Leitfaden_Aufbereitungsaufwand_und_Nachnutzbarkeit_von_Korpora.pdf

Information on data deposit/hosting:

<https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>

DFG standard guidelines for language corpora (in German):

http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standards_sprachkorpora.pdf

Redmine custom workflows:

<https://gitlab.rrz.uni-hamburg.de/hzsk-open-access/redmine-custom-workflows>

CLARIN central monitoring:

<https://monitoring.clarin.eu/>

Documentation and guidelines for repository usage:

<https://github.com/hzsk/hzsk-help-center>

Reviewer Entry

Reviewer 1

Comments:

Accept

An explicit preservation plan that, for example, indicates how and how often the repository checks if files are in need of curation and conversion/migration, and that describes these processes is needed for Compliance Level 4.

Reviewer 2

Comments:

Accept

XI. Data quality

R11. 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:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The HZSK Repository contains mainly spoken language corpora. For projects just starting to compile corpora, the HZSK provides assistance in all areas related to data formats, legal issues and compliance to regulations and requirements given in CLARIN. For projects that are finished with their data collection, the HZSK provides a list of technical and non-technical requirements for data hosting and can supervise a curation process to ensure that data and metadata comply with the repository requirements.

Before ingest, all spoken language corpora hosted at the HZSK Repository have been converted into best practice formats during an external or internal curation process. EXMARaLDA transcription files can also be automatically converted into further best practice formats, e.g. TEI, EAF and, depending on the original transcription guidelines used, CHAT. Corpus metadata and metadata on communications (sessions), speakers and all the relationships and features of physical files is stored in the XML-based EXMARaLDA metadata format Coma, which is also part of the EXMARaLDA system. All main corpus files are stored in open, standard or best practice formats.

For ingest, metadata according to CMDI profiles compliant with the requirements of the HZSK Repository is required. The HZSK provides several modular CMDI profiles that could be complemented by further profiles should project specific description needs make this necessary. The specification of the exact metadata schema is agreed upon between the depositor and the HZSK. The metadata can be created directly in the CMDI format or through conversion from the EXMARaLDA Coma format, or from other formats as agreed upon. The created or converted metadata is then automatically and manually checked for completeness and quality. Since the data deposited with the HZSK Repository has usually been created in external projects, the validation process includes making sure that the resource is comprehensively described through the metadata. In cooperation with the HZSK, the depositor also creates a text description for the resource's landing page. We also ask for comprehensive documentation of the corpus creation and any conventions and schemata used, and encourage depositors to provide further references to work describing the resource and its creation in detail.

To standardize the workflows for data deposit, curation and ingest, the HZSK relies on the Redmine project management software in combination with Git version control. In Redmine, so-called customized trackers have been defined, comprising the individual steps required for each workflow. The specific subtasks can then be assigned and the corrections performed collaboratively by various staff members using the Git version control integration. Within Redmine, there is further, more detailed documentation on these standardized workflows. For automatized corpus data validation and curation, a comprehensive framework (HZSK Corpus Services) is used and continuously extended with new features, making it possible to give data depositors a detailed diagnosis regarding e.g. data consistency problems preventing deposit. Some of these problems can even be fixed automatically. Further details are also given in R7.

The HZSK Repository is continuously gathering information on relevant publications to extend the bibliography on work related to hosted resources. All hosted resources are described in a comprehensive way with references to publications describing them more in detail. Where available, further relevant publications from the bibliography are listed for reference.

The HZSK Repository encourages users to contact its help desk with any questions regarding the resources or tools needed to use them. The HZSK Repository is also integrated into CLARIN, which implements several channels through which members of the designated communities can give feedback on data and metadata hosted by its certified centres.

The CLARIN-D Help Desk, maintained by the HZSK, manages support and feedback workflows for national centres and

various international services, such as the CLARIN VLO. Depending on the type of feedback, help desk agents can thus both forward issues directly to the responsible CLARIN centre and, for issues with a wider impact, contact relevant institutions and bodies at the European level, such as the CLARIN Metadata Curation Task Force, which is responsible for improving and harmonising metadata within the infrastructure.

Furthermore, the discipline-specific working groups within the CLARIN-D project (<https://www.clarin-d.net/en/disciplines>) are yet another communication channel through which the various designated communities can provide more general input and feedback on data and metadata to ensure CLARIN-D centres provide relevant resources and resource descriptions.

References:

Requirements for hosting: <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XII. Workflows

R12. Archiving takes place according to defined workflows from ingest to dissemination.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
3 – The repository is in the implementation phase

Response:

Several workflows are relevant from deposit to ingest and dissemination, and they are documented in various ways and supported by various software solutions. The workflow for data deposit is highly individual since the resources and the contexts from which they stem are rather heterogeneous. Basically, it contains making sure that the technical and non-technical requirements for deposit can be met while obtaining as much relevant information and documentation about the resource as possible and of course setting up a contract as described in R7 and R2 respectively. The cost-benefit analysis guiding decisions on whether to perform curation of the resource is based on a publicly available guideline (see R8 and R11). If non-technical requirements are met, the data and metadata is automatically and manually checked for correctness and completeness using a customized validation framework (cf. R7). Only after necessary curation has taken place and the technical and quality requirements are met, data can be deposited. We are in the process of further extending and standardizing this workflow for various resources types and support it by a customized tracker within our project management software Redmine.

As described in R7, the deposited data is further prepared for ingest. This includes conversion into further standardized formats and formats for web dissemination and sometimes minor corrections or completions of the data. We are currently developing internal workflows based on modern software engineering practices and systems, i.e. Git and Redmine, and model two workflows for this phase based on the resource type one for spoken corpora (of recordings and transcripts) and one for text corpora. The workflows are modelled as customized trackers in Redmine containing all relevant tasks to perform. For tasks implying changes or additions to the resource, individual tasks can be connected to revisions of the data set to ensure transparency.

The workflow for ingest of finalized resources into the HZSK Repository is guided by another customized Redmine tracker and the Fedora Commons system's internal workflows. We also treat our workflows as projects and document changes to them in Redmine.

The workflows for providing access to resources are supported by web forms integrated into the HZSK Repository and the ticketing system of the CLARIN-D Help Desk, and documented as flow charts for internal use.

References:

Communication to depositors about handling of data: <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>

A user's perspective on corpus request workflow is described in the Help pages:

<https://corpora.uni-hamburg.de/help/?lang=en>

Slides from the talk "Corpus Workflows: Formalization and Operationalization of Workflows for the Creation and Publication of Language Corpora" describing workflows used at the HZSK and in the associated project INEL:

https://inel.corpora.uni-hamburg.de/wp-content/uploads/jettka_hedeland-2018-HZSK_INEL_Workflows.pdf

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XIII. Data discovery and identification

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

All CLARIN centres (<https://www.clarin.eu/content/overview-clarin-centres>) provide their metadata in the CMDI format. The Component MetaData Infrastructure (CMDI) (<https://www.clarin.eu/content/component-metadata>) was initiated by CLARIN to provide a flexible framework for describing metadata based on components and concepts. Each metadata record is based on a profile that is registered in the CLARIN CMDI Component Registry (<https://catalog.clarin.eu/ds/ComponentRegistry>). Profiles can make use of components. Those building blocks are also registered in the CMDI Component Registry and describe specific aspects or properties of a resource. Elements of CMDI records link to concept definitions that are stored in external registries (like the CLARIN Concept Registry, <https://openskos.meertens.knaw.nl/ccr/browser/>). Since different communities use different names for the same concepts, linking CMDI elements to concepts enables communities to stick to their terminology while enabling users to find concepts independent of the naming.

A strict requirement for CLARIN centres is to make their metadata available through the established and well documented Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (<https://www.openarchives.org/pmh/>). This standard enables harvesting of the metadata from the repository via http(s).

The CLARIN Virtual Language Observatory (VLO) (<https://vlo.clarin.eu>) harvests the metadata in CMDI format from all CLARIN centres via OAI-PMH. Metadata from all CLARIN centres (and other relevant archives and repositories) are browsable and searchable via the VLO website. CLARIN has defined a set of facets to narrow down the selection of resources in the VLO. These facets are again based on concept sets and allow access to potential heterogeneous metadata stocks. The search in the VLO combines a full text query with a selection of (multiple) values in facets.

Apart from the CLARIN VLO, the HZSK Repository is also registered with e.g. re3data.org and the Bielefeld Academic Search Engine (BASE). Furthermore, through a cooperation with the Fachinformationsdienst Nordeuropa at the Kiel University Library, a subset of the resources of the HZSK Repository have been made available via regular library catalogues, and the feasibility of including this channel on a regular basis for relevant resources is currently being evaluated.

The usage of PIDs is mandatory for resources in CLARIN, thus all resources added to the repository can be referenced using PIDs. For this purpose, the HZSK Repository makes use of a common CLARIN PID service based on the handle system (cf. R0).

Recommended citation formats are provided for all resources and proper attribution one aspect of the general Terms of Use. The HZSK Repository encourages citation using a standardized format for the resource itself, including its PID, but in most cases also provide a second reference publication for the resource, since citation of traditional publications will still likely be more important to data depositors/rights holders.

References:

The CLARIN Virtual Language Observatory: <https://vlo.clarin.eu/>

The HZSK at [re3data.org](https://www.re3data.org/repository/r3d100010139): <https://www.re3data.org/repository/r3d100010139>

The Bielefeld Academic Search Engine (BASE): <https://www.base-search.net/>

Fachinformationsdienst Nordeuropa: <https://www.ub.uni-kiel.de/fach/sondersammlung>

Example of library catalogue corpus record:

<https://www.worldcat.org/title/hamburg-corpus-of-old-swedish-with-syntactic-annotations-hacossa>

CLARIN PID-Service description: <https://www.clarin.eu/sites/default/files/pid-CLARIN-ShortGuide.pdf>

The Handle system: <http://www.handle.net/>

Handle System Implementation: <http://handle.gwdg.de:8080/pidservice/>

Division of labour between CLARIN-D centres:

<https://www.clarin-d.net/en/about/centres/division-of-labour-of-clarin-d-centres>

CLARIN-D's signed co-operation agreement (in German):

<https://www.clarin-d.net/de/ueber/zentren/zusammenarbeit>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

XIV. Data reuse

R14. 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:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

As further outlined in R8, R11 and R13 the HZSK does not allow corpora to be deposited into the repository without successful automatic and manual validation of data and metadata. In preparation for deposit, metadata according to the CMDI metadata profiles used for the data in the HZSK repository is collected. This metadata can be created using EXMARaLDA's Corpus Manager tool (Coma), which is then converted into CMDI and DC metadata. For reference, the

HZSK provides information on useful metadata to include and how to encode it when using the EXMARaLDA Corpus Manager tool for metadata creation and management, but these guidelines are not required to be followed by projects depositing their data. While following the core metadata set guidelines will enable a smooth conversion and more complete integration of the data, apart from general information on the deposited data, the projects can decide on how much information they want to convert into the CMDI format, which is publicly available. It is also possible to provide the metadata directly in the CMDI format or, depending on the specific project context and agreements between the depositor and the HZSK, in other automatically processable formats from which the required CMDI metadata can be generated. In any case, apart from Dublin Core catalogue metadata, content-oriented metadata in the CMDI format for the specific resource type (e.g. spoken corpus or text corpus) must be extractable from the data and metadata provided before it can be deposited. The HZSK uses various CMDI profiles depending on the resource type (cf. <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>), but all profiles share the basic general info component containing the metadata needed for integration of the data into the HZSK and CLARIN infrastructure.

As outlined in R8, all spoken language corpora hosted at the repository have been converted into best practice formats according to the recommendations for standards and tools for compiling language corpora by the Deutsche Forschungsgemeinschaft (DFG, German Research Council). The XML-based EXMARaLDA format is a wide spread format for transcriptions with (mainly manually created) annotations and can be automatically converted into further best practice formats, e.g. EAF, Praat, CHAT. Conversely, as these formats are compatible with the EXMARaLDA formats, other standardized best practice formats can be used for the transcription data, e.g. to better suit increased data complexity etc. since the EXMARaLDA formats cannot explicitly model all features of other relevant formats. We ensure compliance with the EXMARaLDA formats through an internal validation framework, for other best practice formats, corresponding validation of the data is required. The transcription data is however not only provided in the tool formats, but also as automatically generated transcript visualizations of source files to enable qualitative analysis of transcription data, which is a common approach within the designated communities.

To ensure future usability of the data, we only accept data in open well-documented formats (cf. <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>). We convert all data to the EXMARaLDA formats to minimize the migration costs should a complete conversion become necessary. We assume that a conversion of all data into the ISO/TEI standard for Transcription of Spoken Language will be required within the next ten years. The HZSK was involved in the creation of the ISO/TEI standard for Transcription of Spoken Language, which is based on the relevant de facto standard tool formats for transcription used in the designated communities. Though a migration and the use of the ISO/TEI format across the entire HZSK infrastructure is not planned for the very near future, the HZSK is involved in several activities to disseminate this standard and to increase its interoperability with especially the more complex existing (tool) formats that cannot be reliably modelled by the EXMARaLDA formats. Work is also being carried out to ensure interoperability with further platforms and frameworks relevant for the designated communities. The ISO/TEI standard for Transcription of Spoken Language is thus currently viewed as a successor to the various tool based formats in use today and the HZSK Repository and the rest of the HZSK technical infrastructure is being developed to accommodate to this change in the future.

References:

Preferred data and metadata formats of the HZSK Repository: <https://corpora.uni-hamburg.de/hzsk/en/corpus-hosting>

The EXMARaLDA format: <http://www.exmaralda.org/>

The corpora hosted at the HZSK <https://corpora.uni-hamburg.de/repository>

Guideline for the cost-benefit analysis for corpora curation (in German):

https://corpora.uni-hamburg.de/pdf/Leitfaden_Aufbereitungsaufwand_und_Nachnutzbarkeit_von_Korpora.pdf

CMDI component metadata: <http://www.clarin.eu/cmdl>

EXMARaLDA Corpus Manager: <http://exmaralda.org/en/corpus-manager-en/>

ISO Standard "Transcription of Spoken Language": <https://www.iso.org/standard/37338.html>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

TECHNOLOGY

XV. Technical infrastructure

R15. 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:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The repository complies with the OAIS reference model tasks and functions. The repository core builds on the combination of the software Fedora Commons and Islandora, both of which implement the concepts, terminology, and best practices of OAIS. Fedora uses the well-defined Fedora Object XML (FOXML) for content packaging and storage, which makes it possible to recover all stored information independently of the software. It is possible to ingest and disseminate Submission Information Packages (SIPS) and Dissemination Information Packages (DIPS) in standard container formats. The repository architecture generally follows a modular approach that allows for the substitution of individual components. The data consumer has direct access to the archived objects via the web, provided that access requirements have been met.

As a part of CLARIN-D, we are committed to play an active role in the development of CLARIN's repository infrastructure. General plans for maintaining and further developing the infrastructure have been formulated as part of the project proposal. The central goal is to improve the usability of the research infrastructure for typical research tasks such as the retrieval of resources, the evaluation of data or the publication of results. To achieve this, modifications and extensions to a variety of infrastructure components in the repository and in the central infrastructure are necessary. Meetings of all centres to monitor advances in infrastructure development take place quarterly.

Further important goals of infrastructure development are (<https://www.clarin.eu/content/clarin-technology-introduction>):

- To ensure resilience, integrity, and availability of the sustainable repositories and the central infrastructure
- To integrate new resources and tools based on the needs of the user communities
- To allow for better interoperability of tools and resources in the infrastructure
- To enhance the central content search to be more useful in actual research tasks
- To optimize metadata of the resources provided and to enhance user experience in central metadata search

Additional strategic infrastructure planning takes place on the European level in the coordinating committee of the technical centres of the CLARIN ERIC where CLARIN-D also participates.

The guest operating system on the virtual servers provided by RRZ is based on standard operating systems setup. In particular, we use the Ubuntu Long-Term Support version provided. The repository software stack is further described in documentation and articles referenced from <https://corpora.uni-hamburg.de/hzsk/en/about-hzsk-repository>.

References:

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>

Fedora Commons: <http://fedora-commons.org/>

Operating system and software on repository server: (in German)

<https://www.rrz.uni-hamburg.de/services/weitere/virtuelle-server/betriebssysteme.html>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

XVI. Security

R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

We have identified and evaluated the following security risks: data loss or corruption, server downtime, unauthorized access to sensitive data. Risk of data corruption by hardware failure is low. For the event of data corruption we have multiple backup, audit and recovery paths. The repository runs on a virtual machine hosted by the Computing Centre of the Hamburg University (Regionales Rechenzentrum, RRZ). Daily incremental backups allow for a restoration of the whole server within a duration of 24 hours. Bitstream-conservation is provided by the Computing Centre of the Hamburg

University, for a timespan not less than 10 years. The security of physical server and the redundant storage system is managed by the RRZ, relying on their advanced routines to ensure the security and stability of the complex technical infrastructure of the University of Hamburg.

The repository software keeps checksums of repository data streams to prevent data corruption or tampering. Furthermore, there are audit trails and for locally curated data we also maintain the corpus data in a separate version control system on a different virtual machine.

The secure authentication and authorization for restricted material is maintained by way of Shibboleth single sign-on infrastructure using existing institutional accounts or accounts provided by CLARIN. The use of Shibboleth is also a part of the CLARIN infrastructure and a requirement for a CLARIN B centre. The technological implementation is further detailed in Pirinen et al. (2017). To protect the users of the HZSK and its services we have a privacy policy in place and process personal data according to the GÉANT Data Protection Code of Conduct.

For downtime management we utilize multiple online services, such as uptimerobot and statuscake as well as locally installed scripts ran at regular intervals to ensure stability of repository services. In event of server downtime, the repository maintenance staff is notified within 5 minutes.

References:

Information on virtualization at the Computing Centre (RRZ) of the University of Hamburg (in German):
<https://www.rrz.uni-hamburg.de/services/virtuelle-server/daten/virtuelle-server-ordnung-rrz220307.pdf>

Information on backup at the Computing Centre (RRZ) of the University of Hamburg (in German):
<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup.html>

Backup-Policy of the Computing Centre (RRZ) of the University of Hamburg (in German):
<https://www.rrz.uni-hamburg.de/services/datenhaltung/backup/daten/tsm-nutzungsbedingungen.pdf>

Information on IT Security of the Computing Centre (RRZ) of the University of Hamburg (in German):
<https://www.rrz.uni-hamburg.de/services/sicherheit.html>

Pirinen et al. (2017): https://www.clarin.eu/sites/default/files/Pirinen-Jettka-Hedeland-CLARIN2017_paper_22.pdf

Privacy policy: <https://corpora.uni-hamburg.de/hzsk/en/privacy-policy>

GÉANT Data Protection Code of Conduct:
<https://geant3plus.archive.geant.net/uri/dataprotection-code-of-conduct/v1/Pages/default.aspx>

Reviewer Entry

Reviewer 1

Comments:
Accept

It is recommended that during your recertification, you also refer to an explicit disaster plan.

Reviewer 2

Comments:
Accept

APPLICANT FEEDBACK

Comments/feedback

These requirements are not seen as final, and we value your input to improve the core certification procedure. To this end, please leave any comments you wish to make on both the quality of the Catalogue and its relevance to your organization, as well as any other related thoughts.

Response:

Thank you for your work.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments: