



Assessment Information

[CoreTrustSeal Requirements 2017–2019](#)

Repository: CLARINO Bergen Centre
Website: <http://clarino.uib.no>
Certification Date: 18 November 2019

This repository is owned by: **University of Bergen**



CLARINO Bergen Centre

Notes Before Completing the Application

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

True

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

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

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Brief Description of Repository

The CLARINO Bergen Centre is one of four centres in the Norwegian national CLARINO research infrastructure, which in turn is Norway's contribution to the construction and operation of CLARIN ERIC [1]. The aim of the CLARINO Bergen Centre is in line with that of CLARIN ERIC, i.e. to make existing and future language resources easily accessible for researchers, and to bring eScience to humanities disciplines. Thus the CLARINO Bergen Centre is restricted to the domain of language data and tools; datasets may be any collections of language expressions (e.g. corpora, lexicons, termbases) or data about language (e.g. quantitative data obtained from field work or experiments).

The CLARINO Bergen Centre is operated by the University of Bergen (UiB) which is one of the partners in the CLARINO consortium carrying out the CLARINO project funded by the Research Council of Norway (RCN).

The CLARINO Bergen Centre runs a repository (henceforth called the CLARINO Bergen Repository or for short the Repository [2]) offering download and upload of resources. Furthermore the CLARINO Bergen Centre also provides access to deposited resources for online search through the following two online systems for exploration: Corpuscle, a corpus exploration tool, and INESS, the Norwegian Infrastructure for the Exploration of Syntax and Semantics. The CLARINO Bergen Centre also provides COMEDI, an online component metadata editor [3]. All data and services offered by the CLARINO Bergen Centre (the Repository, Corpuscle, INESS and COMEDI) adhere to CLARIN principles for metadata, PIDs, licensing and access.

[1] CLARIN ERIC, the European Research Infrastructure for Language Resources and Technology: <https://www.clarin.eu/>

[2] CLARINO Bergen Repository: <https://repo.clarino.uib.no> (http: will redirect to https:).

[3] The additional online services Corpuscle, INESS and COMEDI are accessible through a top line menu at the main site of the Repository.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Brief Description of the Repository's Designated Community.

The CLARINO Bergen Centre helps researchers working in the Humanities and the Cultural and Social Sciences who have a need for managing, preparing, storing, accessing or analyzing research data related to language. Typical producers of language data are researchers in corpus linguistics, translation, grammar, terminology, lexicography, sociolinguistics, psycholinguistics, philology, philosophy, historical linguistics, political sciences, anthropology, etc. Producers of language processing tools are researchers in natural language processing. Consumers of language data and tools in the CLARINO Bergen Centre are often the same groups as producers; they are interested in analyzing language data and using text processing tools available in the CLARIN infrastructure. Many resources in the CLARINO Bergen Centre originate from Norway, but the centre is open to all CLARIN members, contingent upon licenses.

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

The Repository performs basic curation of the submissions by checking and editing the metadata. The Repository uses the DSpace submission workflow which foresees several curation steps before completing a submission. Curation of submitted items is performed which implies assessment and revision of metadata and quality checks by experts affiliated with the CLARINO Bergen Centre. See checklist at R4. Automatic tools are helping the editors to verify and validate metadata and the integrity of the submitted data [1][2].

The basic curation process foresees the possibility of returning the submission to the submitter for additional changes

before the dataset is published in the repository.

On some occasions the repository assists in enhanced curation, for example, conversion to a different format, enhancement of metadata, or customization of licenses.

[1] CLARIN DSpace metadata info: <https://github.com/ufal/clarin-dspace/wiki/Metadata-info>

[2] DSpace curation system: <https://wiki.duraspace.org/display/DSDOC5x/Curation+System>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

Outsource Partners. If applicable, please list them.

No operational services are outsourced outside of the University of Bergen. The Repository is maintained at the University of Bergen Library, in cooperation with the IT-department and the Department of Linguistic, Literary and Aesthetic Studies at the University of Bergen. See R5 for more details.

The University of Bergen Library is responsible for software maintenance of the repository, implementation of the repository policies (see below) and user support, and commits 2.5 person months per year to these tasks. The Department of Linguistic, Literary and Aesthetic Studies has the scientific coordinator and additional support and commits 3 person months per year to these tasks. The IT department provides the server and basic IT support in terms of back-up, security, network etc.

The IT-department is a unit within the UiB, organised as a part of the central administration. The UiB Library is a unit organised at the level of the faculties and the University Museum. However, the University Library, like the IT-department and the rest of the Central Administration, reports directly to the University Director. A chart is available at [1].

The roles and responsibilities of the IT-department relative to the other departments ("system owners") are defined in the University of Bergen statutes [2]. This obviates the need for Service Level Agreements with each service. See R15 for more details.

Further development of CLARIN DSpace software, of which the Repository is a fork, is performed by a strong community of developers in more than 14 CLARIN centres where features, updates and security measures are presented and solved [3].

[1] UiB organisational chart: <https://www.uib.no/en/about/organisationchart>

[2] Statutes (most relevant parts are in Norwegian only):

https://regler.app.uib.no/regler_en/Part-4-Estate-Economy-and-ICT

[3] CLARIN DSpace development: <https://github.com/ufal/clarin-dspace>

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept

Other Relevant Information.

The current application is part of the re-certification of the CLARINO Bergen Centre as a CLARIN type B centre.

CLARIN will pay the fee for the current CoreTrustSeal certification.

The CLARINO Bergen Centre was certified as a CLARIN type B centre in 2014 [1], which confirms that resources are managed and made accessible through CLARIN compliant interfaces in a stable and persistent way. The CLARINO Bergen Repository was certified with the Data Seal of Approval on 03/11/2015 [2].

[1] CLARIN B centre certificate: <http://hdl.handle.net/11372/DOC-116>

[2] DSA: Our seal and assessment is in the folder for "2014-2017" in the folder "Seals_all" in this public dataset:

<https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:116038/tab/2>

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 mission of the CLARINO Bergen Centre is to provide reliable archiving and documentation as well as easy access to language-related data, tools, services and associated metadata for research purposes (e.g. corpora, lexicons, termbanks, audio and video recordings, grammars, language models, parsers, format converters, lexicon extraction tools, etc.) to scholars in the fields of Social Sciences and Humanities and beyond.

CLARINO implements the Norwegian contribution to CLARIN ERIC [1] which has the objective of advancing research in humanities and social sciences by giving researchers unified single sign-on cross-border access to a platform which integrates language-based resources and advanced tools. This is implemented by the construction and operation of a shared distributed infrastructure that aims at making language resources, technology and expertise available to the humanities and social sciences research communities at large [2].

This work has national approval through the national infrastructure project CLARINO funded by the RCN (2012-2019, grant number 208375) [3]. In turn, Norway's national contribution receives yearly approval by the CLARIN General Assembly. By accepting the RCN grant, UiB has approved of the work plan and mission statement.

The mission statement is shown on the Repository website [4]

[1] CLARIN ERIC: <https://www.clarin.eu>

[2] CLARIN Short Guide: <http://www.clarin.eu/files/centres-CLARIN-ShortGuide.pdf>

[3] The CLARINO project and its funding by the RCN are mentioned in the following RCN documents:

<https://www.forskningradet.no/siteassets/publikasjoner/1254023795613.pdf> and

<https://www.forskningradet.no/siteassets/publikasjoner/1253995988917.pdf>

[4] Mission statement and policies: <https://repo.clarino.uib.no/xmlui/page/about#mission-statement>

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:

Each resource in the CLARINO Bergen Centre is made available under the conditions attached to it. The relations between the involved parties are regulated as follows.

1) For every deposit, a standard contract is made between the CLARINO Bergen Centre and the submitter, the so-called "Distribution License Agreement" [1], in which the rights and duties of each party are described. This agreement acknowledges that the submitter has the right to distribute the resource and gives the CLARINO Bergen Centre the non-exclusive right to distribute the data on their behalf under certain conditions. Submitters can choose among standard licenses [2] or set a custom license. Submitters must be reliably authenticated (registered and logged in, see also R4).

2) Users who download or otherwise access the data are bound by the end user license assigned to the item. Restricted licenses require users to be reliably authenticated and to electronically accept the license before access to the resource is authorized. Only files licensed as publicly available can be downloaded without authentication. The Repository keeps a log of accepted licenses. We do not however log individual downloads, neither do we routinely inspect the activity of users as part of our workflow. So far, our trust in the community of authenticated users to respect the licenses they formally have accepted, seems justified. No incidences have been reported to us. It is also beyond the possibility of the Repository to

check all use of the resources, but if possible abuse is signaled to us, we can retrieve which licenses have been accepted by which authenticated users.

3) Users are bound by the Repository's Terms of Service [3] which specify that users must acknowledge the deposited work in their publications, using the work's persistent identifier, its original author(s)/creator(s), and any publisher where applicable. Full items must not be harvested by robots except transiently for full-text indexing or citation analysis. Full items must not be sold commercially unless explicitly granted by the attached license without formal permission of the copyright holders. Furthermore, the user shall agree to observe best practices regarding research ethics. This includes treating colleagues, stakeholders, customers, suppliers and the public respectfully and professionally, taking into account confidentiality when appropriate, respecting cultural differences and having an open and explicit relationship with government, the public, the private sector and other funders.

Users who disrespect the license conditions or Terms of Service and fail to repair their transgression without delay will be excluded from using the services of the CLARINO Bergen Centre.

[1] Distribution License Agreement (also called "Deposition License Agreement"):

<https://repo.clarino.uib.no/xmlui/page/contract>

[2] List of Available Licenses: <https://repo.clarino.uib.no/xmlui/page/licenses>

[3] Terms of Service: <https://repo.clarino.uib.no/xmlui/page/terms-of-service>

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:

As part of the nationally funded CLARINO infrastructure project, started in 2012, the CLARINO Bergen Centre was established as a permanent infrastructural offer. All its services are currently in the operational phase and are intended to be continued indefinitely. The University of Bergen has committed itself to maintain its part of the CLARINO infrastructure for at least ten years since the start, i.e. until March 2022, and until 2030 if the current update proposal is funded. With a view to local organizational persistence, the Repository has been embedded in the University of Bergen Library, a unit with high stability which curates digital data as well as printed publications.

The University of Bergen Library strategy 2016-2022 [2] states that the University Library is a central part of the research infrastructure of the University of Bergen, and that it shall "develop and utilize technology and infrastructure for retrieval and display of literature and research data" (p. 11, our translation). On p. 12 it is said that the Library shall "offer high level expertise for the production and exchange of metadata", and "follow local, national and international guidelines", further that it shall "contribute to the development of guidelines and systems for storage and reuse of research data". The strategy thus comprises the Library's commitment to Clarino and the running of the Repository.

In the more distant future, continuity will also depend on regular updating in order to remain relevant and compatible in CLARIN; to that end the CLARINO consortium applies regularly for funding to the Research Council of Norway, but this is a competitive process. The continuity of the Repository's integration in the CLARIN ecosystem is therefore contingent on Norway's continued membership in CLARIN ERIC as well as on continued national funding of CLARINO.

In the unlikely case that the Government of Norway would cease its membership and would no longer support future updates, the CLARINO Bergen Centre would maintain its service level for ten years, during which the resources in the Repository would be transferred to one or more sister repositories in CLARIN [2]. Technically, such a transfer will be smooth thanks to the fact that many other CLARIN repositories are forks of the same CLARIN DSpace system. As far as licences are concerned, many resources in the Repository are already cleared for distribution by any CLARIN centre (and in fact some are already duplicated at other CLARIN centres), while some other Distribution License Agreements would have to be renegotiated.

[1] University of Bergen Library strategy 2016-2022 (in Norwegian):

https://www.uib.no/sites/w3.uib.no/files/attachments/ub_strategiplan_endelig_versjon_2016-2022_v4_web.pdf

[2] Preservation policy statement: <https://repo.clarino.uib.no/xmlui/page/about#preservation-policy>

The UiB Library director has in an official letter approved of this policy and confirmed that the Repository holdings will be preserved for at least ten years.

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:

During the submission process, the submitter accepts the repository policy called “Distribution License Agreement” [1] in which they confirm their responsibility regarding the submission, in particular regarding the right to grant the distribution right. The deposited item lifecycle is explained on the website [2]. Information about the requirements and handling of disciplinary and ethical issues is available in the text of the Terms of Service [3].

Any submitter must be authenticated through the federation of ID providers (eduGain and CLARIN Service Provider Federation), which ensures a high level of trust. No anonymous upload of items is allowed. In case of problems, Repository editors can retrieve the exact dates and users who have deposited the resources; for restricted resources, also authenticated users who have accessed (downloaded) the resources are logged. This means that, if needed, users who do not comply to the license or the Terms of Service can be blacklisted: they can be denied further access to the repository and the research community may be made aware of any abuse.

Each submission is validated automatically and manually by an expert Repository editor, whose task includes a check if privacy and ethical issues are respected. If the editor detects problems with the submission, he/she will interact with the depositor, and request or suggest modifications. The submission will not be approved until all problems are resolved. Repository editors are data experts who are careful not to make changes which in themselves may create privacy or ethical issues.

This is the manual checklist used by the administrator for each submission:

"On the basis of the submitted data and metadata for new submissions, check that

1. There are no GDPR/personal data issues represented by publishing the datasets in Clarino Bergen Repository. The Repository does not accept data sets containing personal data. Anonymised data sets are generally accepted. If in doubt, check that the submitter has cleared the publication procedure with the Norwegian Centre for Research Data (NSD).

2. There are no IPR/Copyright issues represented by publishing the datasets in Clarino Bergen Repository. Assess the submitter's acknowledgement that (s)he is the holder of the rights to the data. If in doubt, check that depositing rights are cleared with other rightholders that might be implied.

3. Data formats are compliant with CLARIN recommendations (LRT).
If not: Ask the submitter to convert the data. Assist in data conversion if necessary and possible.

4. CMDI metadata file is provided.
If not: Ask the submitter to provide CMDI metadata. Assist in producing metadata if necessary.

Remember for old submissions that any changes to data need to take the form of submitting a new version of the data as a separate data set. Submitted data sets can not be changed."

All Norwegian research producing data with potential privacy issues must clear their procedures with the Norwegian Centre for Research Data NSD [4]. As stated in the checklist above, the Repository does not accept data sets containing overt sensitive personal data. Anonymized (or pseudonymized) data sets are generally accepted. If in doubt, we check that the submitter has cleared the publication procedure with the NSD. Since even anonymized or pseudonymized data may carry residual risks of revealing personal information, such data are normally released under a restricted license (CLARIN RES license with the additional PRIV condition) [5].

All interaction with the Repository is via secure protocols (https).

The CLARIN Legal Information Platform [6], set up and run by the CLARIN Legal and Ethical Issues Committee (CLIC) of CLARIN ERIC, provides relevant advice and organises periodic training sessions in management of data with a disclosure

risks.

[1] Distribution License Agreement: <https://repo.clarino.uib.no/xmlui/page/contract>

[2] Deposited Item Lifecycle: <https://repo.clarino.uib.no/xmlui/page/item-lifecycle>

[3] Terms of service: <https://repo.clarino.uib.no/xmlui/page/about#terms-of-service>

[4] NSD: <http://nsd.no>

[5] CLARIN restricted licenses: <https://www.kielipankki.fi/support/clarin-eula/>

[6] CLARIN Legal Information Platform: <https://www.clarin.eu/content/legal-information-platform>

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 Repository is hosted by UiB, a medium-sized academic institution which is appropriate to the Repository's main audience of academic researchers and students. The University of Bergen is a member of the CLARINO consortium, which in turn is implementing Norway's contribution to the CLARIN ERIC. The CLARINO Bergen Centre is a certified

centre type B in the CLARIN ERIC [1], the governance of which strongly contributes towards guiding the CLARINO Bergen Centre in carrying out its mission.

UiB has committed itself to maintain its part of the CLARINO infrastructure for at least ten years, i.e. until March 2022, and until 2030 if the current update proposal is funded.

The UiB Library, which is directly responsible for the operation of the Repository, is a highly stable unit with a long term mission that includes the curation of digital data. The UiB Library is committed to CLARINO with 2,5 person months pr. year. IT-resources for the Repository are provided by the University's IT Department (see R0).

The Repository is located at the Library's Section for Education and Research Support, and run by the Digital Development team there in cooperation with the Department of Linguistic, Literary and Aesthetic Studies. At the Digital Development team the following persons are assigned to CLARINO tasks: Chief Engineers Hemed Ali Al Ruwehy and Øyvind Liland Gjesdal, Senior Academic Librarian Rune Kyrkjebø and Senior Executive Officer Paul Meurer. Al Ruwehy is responsible for technical maintenance, upgrades and development. He keeps in touch with LINDAT/CLARIN in Prague, which is the leading technical centre among the CLARIN centres running DSpace repositories. Al Ruwehy has the main responsibility for co-operation with the University IT-department, and he provides the day-to-day maintenance and technical support for our repository. Liland Gjesdal represents a doubling of this essential function within the team. The two Chief Engineers conduct larger tasks, like software upgrades, together. Meurer maintains the additional services (Corpuscle, INESS and COMEDI) and their interaction with the Repository. Kyrkjebø is assigned by the library to work with CLARINO on the scientific and administration and user support levels. Administration of submissions is mainly handled by Kyrkjebø.

Principal Librarian Tarje Sælen Lavik heads the Digital Development team and coordinates the tasks based on regular status meetings where progress is reported and challenges are discussed. The team belongs to the Section for Education and Research Support, with Karin Cecilia Rydving as Section Head. Sælen Lavik reports to Rydving. Rydving reports to the Library Director Maria-Carme Torras. They answer for the Library's strategy and its long- and shortterm commitments. All that are mentioned so far are permanent full time staff, except Meurer who is 60%. Also readily available is the advice and assistance of Professor Koenraad De Smedt at the Department of Linguistic, Literary and Aesthetic Studies, with whom the library team works closely together. As National Coordinator and leader of the CLARINO project and the Bergen Centre, De Smedt coordinates also the library team, and ensures that the library is always informed about CLARINO and CLARIN activities and possibilities to gain new knowledge, acquire contacts and share experience within the community.

The Repository website has contact information [2].

[1] Certificate for type B status: <http://hdl.handle.net/11372/DOC-116>

This certification is up for renewal as soon as the CoreTrustSeal (which is part of the criteria) is granted.

[2] Contact information: <https://repo.clarino.uib.no/xmlui/page/contact>

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:

The Repository is in practice managed, operated and maintained by members of the Digital Development team at the University of Bergen Library (see R5), in cooperation with the Department of Linguistic, Literary and Aesthetic Studies and with additional advice by the Norwegian School of Economics at Bergen. The members of the Digital Development team are employees selected for their high level knowledge and skills and extensive experience in digital data management. The team is a part of the Section for Education and Research Support. The responsibilities for the Section for Education and Research Support is stated on internal, Norwegian webpages as:

- Guidance and instruction in information literacy and academic writing
- Collection development and curation
- Open science: Digital publications and research data
- Bibliometry
- Copyright

- Development of digital services [the overall task of the Digital Development team]
- Metadata
- Outreach

Some of the main digital resources run by the Digital Development team can be accessed from the link page "Writing and Research" at the University Library homepage [1]. The CLARINO Bergen Centre Repository is directly accessible from that page, as are several resources under "Digital Text Archives", and the "Manuscripts and Photographies" overview which leads on to the MARCUS homepage (direct link here: [2]). MARCUS is a Linked Open Data portal for the University Library's Special Collections, developed and run by the team.

The Digital Development team does not have its own webpage with a mandate. The current University Library Staffing Plan ("Bemanningsplan_endelig_versjon_151117" effectuated from 2018), which is an internal document, states tasks and competencies for the team. Among the tasks listed there, are (our translation):

- Development and running of infrastructure and digital services for teaching, research and outreach.
- Co-operation on infrastructure and digital services for teaching, research and outreach.
- Work on encoded text resources using XML technologies.
- Work on semantic technologies.
- Development of web applications and free software.
- Running of servers [note: always in co-operation with the University IT-department]

The staff regularly participate at scientific meetings at national and international levels where research community members are present, including annual CLARIN conferences. They also participate in training and knowledge sharing activities and CLARIN committee meetings. Al Ruweh participated in the CLARIN workshop on DSpace development in 2016 and remains in contact with the CLARIN DSpace development group. Paul Meurer received the Steven Krauwer Award for CLARIN Achievements in 2017.

Since the Repository runs on servers at the IT-department of the university, qualified personnel at that department provide the necessary expertise for running the servers.

Overall scientific coordination of the CLARINO Bergen Centre is in the hands of Prof. Koenraad De Smedt at the Department of Linguistic, Literary and Aesthetic Studies, who is a member of the CLARIN National Coordinator's Forum and makes sure that community relevance as reflected in CLARIN strategies percolate to that of the Repository.

The Repository staff provides feedback to all requests and submitters through email. In some cases, community members request personal discussions and guidance, in particular in the context of data management plans to be drafted at the start of new research projects. In cases of work-intensive data management, a contract with the depositing researcher of research group is drawn up, which may involve payment.

The coordinator and other staff members provide guidance to users through lectures and training events, for instance, at the National Research School for Linguistics and Philology [3].

[1] <https://www.uib.no/en/ub>

[2] <http://marcus.uib.no/>

[3] An extended training course for the research school was given by Koenraad De Smedt and Gunn Inger Lyse Samdal on May 30 to June 3, 2016. Another course is in preparation.

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 Repository is based on the CLARIN DSpace which is a fork of DSpace software [1]. To verify that a digital object has not been altered or corrupted, the repository uses MD5 checksums for all objects and checks these periodically.

The Repository automatically performs regular checks on the integrity and the file formats of data. There is a list of supported and known formats whose consistency are regularly checked using existing tools (e.g., integrity testing of bzip format is done using `bzip -t`). A report is then sent to the editors and administrators who keep track of all used formats. If a new, more commonly used format emerges, the Repository will add it to the recommendation.

We adhere to the principles regarding standards and data formats are those formulated by CLARIN [2], but if the submitter has a good reason to use a non-open format, we accept that if properly documented. Items in the repository are normally only made web-accessible in the same format as submitted.

All submitters must be authenticated through the federation of ID providers (eduGain and CLARIN Service Provider Federation). Actions regarding each Repository item by submitter and administrators are logged by the DSpace software and added to the DSpace metadata for the object. Examples of logged actions are the uploading of files, editing of metadata, approval of the submission and publishing of the submission. The logs are directly available for administrators.

Once deposited and archived, the submitted data sets can be changed neither by the submitter nor by editors. As stated in the Distribution License Agreement, within the repository no alteration of the submitted data will be made. This ensures that data is authentic and it is also important for the assigned persistent identifiers, which must always refer to the same content. This point is well understood by the administrators (Kyrkjebø, Al Ruwehy, Gjesdal) and is stressed on our internal submissions checklist (see R4). It is also explicated on the About-page (see Preservation Policy) [3]

Only the administrators of the Repository have the rights to make changes; also these actions are logged (see above). We consider making minor changes to the metadata in case of errors; in this case, submitters should contact the help-desk for requesting such changes as described in the "Deposited Item Lifecycle" [4]. For non-trivial changes, including changes to data formats, we require a new submission and will assign a new PID.

[1] DSpace: <https://duraspace.org/dspace/>

[2] CLARIN standards and formats: <https://www.clarin.eu/content/standards-and-formats>

[3] <https://repo.clarino.uib.no/xmlui/page/about#preservation-policy>

[4] Deposited Item Lifecycle: <https://repo.clarino.uib.no/xmlui/page/item-lifecycle>

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 Repository documents well-defined procedures on how to deposit various collection types [1]. The repository mainly supports the following collection types to be submitted: corpora, lexical-conceptual resources, language descriptions, software, tools, web-services, and quantitative data from language surveys and experiments. The addition of other types is possible. Types corresponding to published CLARIN metadata profiles are recommended.

The Repository relies on the group of metadata standards around CMDI (ISO-CD 24622-1) [2]. The CMDI best practices guide promotes the inclusion of sufficient information to interpret the data in the context of long-term preservation [3]. Submitters are encouraged to provide a CMDI record based on a specific profile which is suitable for the submitted data type (e.g. corpus, lexical resource, etc.). In case the submitter does not upload a CMDI file, our administrator will ask for this to be provided. The automatically generated CMDI metadata exported by the DSpace application guarantees findability, but are not as rich in content as a tailored CMDI file normally would be. Part of our service is to offer assistance in generating a richer CMDI file using our tool COMEDI, which is accessible via a button on the Repository homepage, using the same federated login as for the Repository [4].

Each submitter must be authenticated through the federation of ID providers (eduGain and CLARIN Service Provider Federation), which ensures a high level of trust. No anonymous upload of items is allowed. In case of problems, Repository editors can retrieve the exact dates and users who have deposited the resources.

Submitters are referred to our FAQ page [5] which provides the list of data formats recommended by CLARIN [5]. When the submitted items contain attached bitstreams, metadata curators manually verify whether they meet the requirements of integrity, authenticity and availability. Metadata curators are always in contact with the depositor(s) in order to obtain missing information if there is any.

If the format is unknown or not in the list of the recommended standard formats [6], it must be well documented and the documentation must be either part of the submission or the metadata must contain a link to it. The final decision on acceptance/rejection of such submissions is taken by the Repository administrators in collaboration with the CLARINO metadata committee.

If the minimum formal requirements are not met during the submission process, the interface will not allow the user to complete the submission. After submission, the item is reviewed by metadata experts that will check for the quality of the metadata. Editors are also responsible of verifying that the data submitted are in a valid, correct and accepted format. However, as stated in the Distribution License Agreement, submitters are ultimately responsible for the quality of their own data. In case the submission does not comply with our expectations, the submission is returned to the data provider for correction.

Once deposited and archived, the submitted data cannot be changed neither by the submitter nor by editors. In case of severe problems e.g. regarding data quality or licenses, data can be retracted. If a dataset needs to be changed in any way, the submitter will create a new version which will be assigned a new persistent identifier while the old dataset will normally remain in place. This will also be the case if data needs to be migrated to a different format. Any new versions will be regarded as new datasets to which the entire submission procedure applies.

[1] How to Deposit: <https://repo.clarino.uib.no/xmlui/page/deposit>

[2] About CMDI: <https://www.clarin.eu/content/component-metadata>

CMDI schema: http://catalog.clarin.eu/ds/ComponentRegistry/rest/registry/profiles/clarin.eu:cr1:p_1349361150622/xsd

[3] Best practices guide: <https://www.clarin.eu/content/cmd-i-best-practices-guide>

[4] COMEDI: <http://clarino.uib.no/comedi/page>

[5] FAQ: <https://repo.clarino.uib.no/xmlui/page/faq?locale-attribute#what-submissions-do-you-accept>

[6] Standards and formats recommended by CLARIN: <https://www.clarin.eu/content/standards-and-formats>

Metadata Standards: <https://www.clarin.eu/sites/default/files/Standards%20for%20LRT-v6.pdf>

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:

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:

Data storage is done in a DSpace repository application which is part of the UiB IT infrastructure for computing and data storage. For the first step, the ingestion process, Submission Information Packages (SIPs) are received for curating and are assigned to a task pool where our curators can process them. The default ingestion process is done through our web based interface. There is a number of pre-configured supported SIP formats (see [1]). For the second step, the archival storage, one of our curators takes charge of the submission, the metadata are updated (added, deleted, modified), and the submitted bitstreams are validated. In general, the curators ensure consistency and quality of each submission. If a curator approves an item, the Archival Information Package is available.

All backups follow standardized ways of using MD5 checksums for determining consistency. For securing the long-term storage of digital data, the IBM Tivoli Storage Manager (TSM) is used by the IT-department for taking daily backups [2]. The TSM backup includes daily database dumps and the resource directories containing the deposited items. We also have a daily snapshot taken of the installation itself, going back at least 2 weeks. For recovery, the IT department at the University of Bergen keeps its virtualized services including servers in different buildings that are in different locations. In this way, if one cluster fails, the other one will be picked up. Also, there is a plan to move some virtual server services to a cloud infrastructure, where CLARINO Bergen Centre will be one of the beneficiaries. The repository application source code is maintained at the UiB Git repository. The UiB IT department runs RedHat virtualization servers where the Repository is hosted. These machines have long term support, are strongly security updated, and comply to ISO/IEC 27001 (see R16). Since the Repository does not store highly sensitive data the security measures are generally considered as sufficiently safe.

The above procedures are internally documented but due to their technical nature, they are not described at the same level of detail in the documentaion on the Repository webpages.

[1] SIP formats: <https://wiki.duraspace.org/display/DSDOC18/Importing+and+Exporting+Content+via+Packages#ImportingandExportingContentviaPackages-SupportedPackageFormats>

[2] TSM: https://www.ibm.com/support/knowledgecenter/ko/SSGSG7_7.1.0/com.ibm.itm.ic.doc/welcome.html

Reviewer Entry

Reviewer 1

Comments:
Accept

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:

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 digital assets of the repository will be long-term preserved as part of the UiB digital assets. UiB has long-term support of its IT infrastructures, where the repository is hosted, and the IT department guarantees persistence of data and regular backup. The repository is covered by University of Bergen IT-department backup routines, security measures and crisis management for all of its systems, complying to ISO/IEC 27001 (see also R9 and R16 for details).

For each item in the Repository, the CLARINO Bergen Centre has a Distribution License Agreement (see also R2). The default agreement is an easy-to-understand text [1] giving the CLARINO Bergen Centre the non-exclusive right to store the item and distribute it under specified conditions. This agreement does not imply transfer of ownership. The agreement does not by default specify a time period or limit. Thus the Repository may store and distribute the data for as long as it wants, and retract it if wishes to do so. The agreement gives the Repository the right to move the data to new media as necessary for long-term preservation. The repository also has the right to convert the data to new formats; such conversions (resulting in new items) have until now not been necessary, since the Repository promotes widely accepted

open formats which tend to stand the test of time. Custom distribution license agreements are possible if needed.

Relevant actions during long-term operation are automatically logged by the DSpace system.

In terms of community embedding, the Repository's strong link to CLARIN and its common strategy aimed at open data, open standards and formats, and open software maintained by an international group of CLARIN centres is a major factor promoting longevity of data access. The core of the plan for managing the function of long-term preservation is therefore to maintain this strong link to CLARIN and the community behind this large European research infrastructure. Participation in CLARIN committees and working groups is part of this plan.

In order to realize this data preservation plan, the UiB Library is committed to spend the necessary efforts to sustain long term access to and preservation of the data contained in the CLARINO Bergen Centre Repository (see also R3 and R5).

The Repository preservation policy is documented at [2].

[1] Distribution License Agreement (also called "Deposition License Agreement"):

<https://repo.clarino.uib.no/xmlui/page/contract>

[2] <https://repo.clarino.uib.no/xmlui/page/about#preservation-policy>

Reviewer Entry

Reviewer 1

Comments:

Accept

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 CLARINO Bergen Repository has experts who are able to address technical data and metadata quality (see R5 and R6 for details).

Every submitted item undergoes technical assessment of data metadata quality by experts affiliated with the CLARINO Bergen Centre. Automatic tools are helping the editors to verify and validate metadata and the integrity of the submitted data [1][2]. The basic curation process foresees the possibility of returning the submission to the submitter for additional changes before the dataset is published in the repository. On some occasions the repository assists in enhanced curation, for example, conversion to a different format, enhancement of metadata, or customization of licenses.

The metadata policy ensures that sufficient information is available for end users (see R8 for details). It is, however, ultimately up to end users to assess to which extent the data suit their own goals. The Repository does not provide any guarantees regarding data quality or suitability, since this is the responsibility of the rights holders. Neither does the Repository require that sufficient instructions are provided to use certain data for certain purposes.

The CLARINO Bergen Centre provides an online tool called COMEDI [3] for creating, modifying and validating CMDI-conformant metadata [2]. In this way we help users/depositors of our repository and metadata curators to ensure that resource quality is maintained. The use of COMEDI is however not required for all submissions.

The Repository does not provide links to related works. The portal CLARIN Virtual Language Observatory (VLO) [4] harvests the CMDI metadata from the CLARINO Bergen Repository, together with that from other Repositories, and can provide lists of items that match faceted search criteria.

The Repository has contact information for reporting issues. Furthermore, in the VLO overview, i.e. when browsing or searching for relevant resources, and on the individual resource pages displaying further information on a specific resource, the user can report an issue or give feedback on metadata records or resources using a ticketing system. The help desk agents can thus forward issues directly to our Repository for correcting and/or harmonizing metadata if needed.

Data and metadata quality is also promoted through guidance to users through lectures and training events [5]. We also encourage users to attend CLARIN training courses and other user involvement actions [6].

[1] CLARIN DSpace metadata info: <https://github.com/ufal/clarin-dspace/wiki/Metadata-info>

[2] DSpace curation system: <https://wiki.duraspace.org/display/DSDOC5x/Curation+System>

[3] COMEDI: <http://clarino.uib.no/comedi>

[4] VLO: <https://vlo.clarin.eu>

[5] An extended training course for the research school was given by Koenraad De Smedt and Gunn Inger Lyse Samdal on May 30 to June 3, 2016. Another course is in preparation.

[6] CLARIN user involvement actions: <https://www.clarin.eu/content/clarin-for-researchers>

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:

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 has a well-defined and documented workflow for the curation process and a procedural documentation for archiving of data [1]. The repository also documents deposit life cycle for creating, editing or deleting resources [2]. The workflow is managed by Rune Kyrkjebø and his colleagues at the Section for Digital Development of the UiB Library (see also R5 for more detail on the organizational infrastructure).

Ingest includes appraisal and selection of data by the Repository editors in accordance with the mission of the repository [3]. In certain cases, requests for the deposit of valuable but unsuitable data may be deflected to other repositories.

The Repository distinguishes between known submitters and unknown ones, where submissions from the latter will be specially validated and verified. In the day-to-day running of the repository, the University of Bergen Library relies upon linguistic advice from CLARINO partners when needed in connection with metadata or licensing questions. The necessary skills for the workflow to function is provided on a long term basis.

The workflow includes checks for sensitive data (e.g. containing personal identifying information) since the Repository does not accept such data (see also R4). Sufficient levels of security are in place for handling non-sensitive data in the workflow (see also R9).

Submissions that do not meet all criteria, e.g. due to inconsistent formats, incomplete or invalid metadata, missing or inappropriate license, personal data issues, etc. are rejected.

[1] Deposit procedure: <https://repo.clarino.uib.no/xmlui/page/deposit>

[2] Item Lifecycle: <https://repo.clarino.uib.no/xmlui/page/item-lifecycle>

[3] Mission statement: <https://repo.clarino.uib.no/xmlui/page/about#mission-statement>

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:

The items in the Repositories are findable in several ways. The Repository is browsable and searchable with keywords and facets. Metadata on all resources in CLARINO Bergen Centre can be harvested by authorized systems through OAI-PMH endpoint. They are regularly harvested and findable through CLARIN Virtual Language Observatory (VLO) which supports keywords and faceted search [1]. They are also harvested into a national catalog maintained by the National Library of Norway and can be discovered there [2].

An example of a record resulting from a search in the VLO and describing an item in the Repository is the following:

https://vlo.clarin.eu/record?4&count=3&docId=http_58__47__47_hdl.handle.net_47_11509_47_124_64_format_61_cmdi

This record contains a handle referring to the landing page in the Repository.

We assign persistent Identifiers (PIDs) in the form of handles to all items by the repository allow for reliable citation. The Repository generates full bibliographical references on the fly, both in textual format (author/year style) and in BibTex.

Access to data is subject to registration and authentication of the user, with the exception of non-restricted, public data.

Submitters of data in a non-recommended format will be encouraged to convert the data to a recommended format.

Administrators will ask for this to be done before approval of the submission. Our staff will assist with advice on conversion of data. If time and resources allow it, we might offer to assist with the data conversion itself. This is stated in our internal submission checklist and on the Deposited Item Lifecycle-page [3].

Resources are provided by the community and to the community in commonly used formats. As stated in our FAQ, "We accept any linguistic and/or NLP data and tools: corpora, treebanks, lexica, but also trained language models, parsers, taggers, MT systems, linguistic web services, etc." [4].

[1] VLO: <https://vlo.clarin.eu/>

[2] Catalog at the National Library of Norway:

<https://www.nb.no/sprakbanken/#ticketsfrom?lang=en&query=alle&tokens=&from=1&size=12&collection=clarino>

[3] Deposited Item Lifecycle (under Edited Item): <https://repo.clarino.uib.no/xmlui/page/item-lifecycle>

[4] FAQ: <https://repo.clarino.uib.no/xmlui/page/faq#what-submissions-do-you-accept>

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:

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 CLARINO Bergen Centre has been established with data reusability as one of the main components in mind. It provides clear information about what metadata are required and how the data and metadata are disseminated [1]. Depositors can fill metadata based on the profiles for specific resource types so that appropriate information is given to support the understanding and use of the data. The Repository does, however, not provide any guarantees for suitability of the data for any purposes of reuse.

The Repository metadata strategy is based on the CLARIN metadata strategy which supports the understanding and reuse of the data (see R8 for details). All metadata comply with CMDI profiles/schemas in the Component Registry [2] and all of them map to Dublin Core. The repository supports several protocols for data sharing and reusability such as OAI-PMH, OAI-ORE and offers resources to the community in different metadata formats such as CMDI, OAI_DC, HTML, BIBTEX, ORE etc. [3].

The strong embedding of the Repository activities within the CLARIN community, including participation in the CLARIN CMDI Taskforce and use of the CMDI component registry, promotes longevity of the metadata strategy, ensuring the validity of the metadata over time and allowing for further developments, e.g. the future migration from CMDI 1.1 to CMDI 1.2. Since this migration has not yet been completed for the Repository, the compliance level is set at 3 (implementation phase) even if the Repository is fully operational [4].

The Repository is open to adoption of future standards and formats and to further development of the CLARIN DSpace system thanks to cooperation with a strong community of developers and users all over the CLARIN centres in Europe. This makes it easier to incorporate any features based on future requirements. As an example, we mention that CMDI profiles were originally not supported in DSpace out of the box, but were later incorporated in the system by the CLARIN developer community. The Repository supports submission of new versions of the resources in case there is a need for enriching or reformatting submitted datasets.

The Repository itself has not yet given any training to researchers in data reuse. We are, however, aware of the fact that several academic courses and researcher training courses have provided instruction and demonstration of reuse of data in the Repository.

[1] Metadata: <https://repo.clarino.uib.no/xmlui/page/metadata>

[2] CLARIN Component Registry:

https://catalog.clarin.eu/ds/ComponentRegistry?registrySpace=published&itemId=clarin.eu:cr1:p_1403526079380

[3] List of metadata formats: <https://repo.clarino.uib.no/oai/request?verb=ListMetadataFormats>

[4] Adoption of CMDI 1.2 is foreseen in the proposed CLARINO+ project which has been submitted for funding.

Reviewer Entry

Reviewer 1

Comments:

Accept

Reviewer 2

Comments:

Accept.

This is very close to being Compliance Level 4. The outlined developments are expected to be in place by the renewal of your CoreTrustSeal certification.

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 basic middleware for the Repository is CLARIN DSpace [1], which is in use by several repositories in CLARIN member countries, which means that a large European research community is familiar with its purpose, features and interface.

This software is a modification of the generic DSpace [2], a full-fledged turnkey institutional repository platform. It was modified by the LINDAT/CLARIN centre [3] at the Charles University in Prague in order to comply to CLARIN specs and requirements regarding standards, formats, authentication, etc. [4]. The CLARIN DSpace software is currently supported, documented and widely used by a community of CLARIN technical centres who produce regular improved versions, maintain a wiki [5] and organize development workshops [6]. The CLARINO Bergen Centre plans to continue its participation in this further development and will regularly upgrade to improved and tested versions of this software.

At the CLARINO Bergen Repository, this software runs on a virtual server maintained by the IT department of UiB. The virtualization with full backup allows for high portability if necessary. The virtual server hardware is placed in a secure air-conditioned server room with restricted access. Permanent connectivity to the internet is provided in the Norwegian academic network maintained by Uninett, the Norwegian NREN [7], boasting very high capacity, redundancy, transparency, and neutrality, and 100 Gbps connection to GÉANT. The UiB IT department runs RedHat virtualization servers where the Repository is hosted. The machines have long term support, are strongly security updated, and comply to ISO/IEC 27001 (see reference at R16). All communication with the systems is based on secure protocols. Since the Repository does not store highly sensitive data these measures are generally considered as sufficiently safe.

For information on backup and long term storage see R9.

The roles and responsibilities of the IT-department relative to the other departments ("system owners") are defined in the University of Bergen statutes [8]. This obviates the need for Service Level Agreements with each service. See also R15.

[1] CLARIN DSpace <https://github.com/ufal/clarin-dspace>

[2] DSpace: <https://duraspace.org/dspace/>

[3] LINDAT/CLARIN: <https://lindat.mff.cuni.cz/en>

[4] CLARIN standards and formats: <https://www.clarin.eu/content/standards-and-formats>

[5] CLARIN DSpace wiki: <https://github.com/ufal/clarin-dspace/wiki>

[6] CLARIN DSpace workshop: <https://www.clarin.eu/blog/clarin-dspace-workshop>

[7] Uninett: <https://www.uninett.no>

[8] Statutes (most relevant parts are in Norwegian only):

https://regler.app.uib.no/regler_en/Part-4-Estate-Economy-and-ICT

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:

UiB has documented procedures for information security and personal protection [1] consisting of goals, rules and management procedures which also govern the Repository. They are anchored in laws, regulations and management responsibilities as defined in the organization of the university, with the University Board as the highest responsible. All activity must comply with various Norwegian laws regarding personal data protection (including the GDPR), electronic

communication, transparency, archiving, special protection, copyright, higher education, health research, health registers, and health personnel, and must also comply with the general UiB regulations. Confidentiality must be guaranteed in accordance with all these laws and regulations. Personal information must not be processed to a larger extent than necessary and must be removed when no longer needed for defined academic purposes. Personal data should be correct. Everyone who processes personal data must be known with relevant rules and requirements.

The framework of its IT policy is within the standard ISO/IEC 27001: 2013 [2]. This standard is used to identify all domains of security which are applicable to the university's IT systems including those of the Repository. The standard is continually operationalized by the IT-department. The operative responsibility for IT security is with the IT director.

In line with the regulations, the Repository has a documented privacy policy [3] describing its jurisdiction, the personal data processed, the purpose of processing personal data, how to access, rectify and delete personal data, data retention and the data protection code of conduct. The policy also states that no personal data are disclosed to anyone outside of the Repository team. Personal data such as name and email are used to allow user authentication and authorization, to log the signing of licenses, to contact the user if necessary, and for monitoring and technical issue solving. Personal data are also used (in aggregated form only) for statistics on usage by groups in society, as required in reports by the Research Council of Norway.

Secure protocols (mainly https) are used for all communication with the system.

For information on backup procedures, see R9.

The IT department follows an internal action plan that is executed in case of major security breaches. The IT department regularly runs crisis management exercises, for example to quickly remedy the situation should a machine crash occur. In such situations a crisis management group enters into force. There are also established routines for reporting security incidences to the relevant authorities.

[1] UiB security policy (in Norwegian): <https://regler.app.uib.no/regler/Del-4-OEkonomi-eiendom-og-IKT/4.3-Informasjons-og-kommunikasjonsteknologi/Styringssystem-for-informasjonssikkerhet-Del-1-Styrende-del>

[2] UiB security policy, executive part (in Norwegian). The ISO/IEC 27001: 2013 standard is used in the appendix "Oversikt over etablerte sikringstiltak ved UiB " which gives the overview of established IT security measures at the University of Bergen. See: "7. Vedlegg" at: <https://regler.app.uib.no/regler/Del-4-OEkonomi-eiendom-og-IKT/4.3-Informasjons-og-kommunikasjonsteknologi/Styringssystem-for-informasjonssikkerhet-Gjennomfoerende-del>

[3] Privacy policy: <https://repo.clarino.uib.no/xmlui/page/privacypolicy>

Reviewer Entry

Reviewer 1

Comments:

Accept

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:

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments: