Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository Språkbanken CLARIN Repository complies with the guidelines version 2014-2017 set by the Data Seal of Approval Board.
The afore-mentioned repository has therefore acquired the Data Seal of Approval of 2013 on November 11, 2016.

The Trusted Digital repository is allowed to place an image of the Data Seal of Approval logo corresponding to the guidelines version date on their website. This image must link to this file which is hosted on the Data Seal of Approval website.

Yours sincerely,

The Data Seal of Approval Board

Assessment Information

Guidelines Version:	2014-2017 | July 19, 2013
Guidelines Information Booklet:	DSA-booklet_2014-2017.pdf
All Guidelines Documentation:	Documentation
Repository:	Språkbanken CLARIN Repository
Seal Acquiry Date:	Nov. 11, 2016
For the latest version of the awarded DSA for this repository please visit our website:	http://assessment.datasealofapproval.org/seals/
Previously Acquired Seals:	None
This repository is owned by:	**Språkbanken** Sweden T +46317860000 E sb-info@svenska.gu.se W http://spraakbanken.gu.se/

Assessment

0. Repository Context

Applicant Entry

Self-assessment statement:

The Språkbanken CLARIN Repository is a repository for language technology resources and language technology tools available at:

https://spraakbanken.gu.se/

Our repository is part of the Swedish Swe-Clarin project which is the Swedish part of the

Common Language Resources and Technology Infrastructure (CLARIN) project. Språkbanken is the national coordinator of Swe-Clarin.

More about Swe-Clarin can be found here:

https://sweclarin.se/eng

More about CLARIN and the CLARIN European Research Infrastructure Consortium (CLARIN ERIC) can be found here:

https://www.clarin.eu/

Språkbanken is an organizational part of University of Gothenburg and as such staff, equipment, technical infrastructure, as well as rules and policies are to a large extent decided/managed by the University.

The aim of the repository is to provide discoverable language technology resources and language technology tools in order to fulfil the specifications of a CLARIN type B Centre for Språkbanken.

The Språkbanken CLARIN Repository serves as an archive for language technology resources for documenting and processing Swedish since its first written forms appeared and the language techology tools to work with said resource in processing of texts for research over past time and into the future. This includes corpora and lexica compiled by Språkbanken as well as compiled by external resource providers:

https://spraakbanken.gu.se/eng/resources/

The same holds for language technology tools for research developed by Språkbanken as well as building on tools developed by others:

https://spraakbanken.gu.se/eng/research/

The repository software we use is the LINDAT/CLARIN modified DSpace. This means our repository provides persistent Identifiers, authorisation and authentication, and sharing of metadata and data. Data harvesting according to OAI-PMH is available.

The repository is open for self-deposit by users, by a documented procedure:

https://repo.spraakbanken.gu.se/xmlui/page/deposit

Submissions will be reviewed by Språkbanken staff at University of Gothenburg.

In compliance with CLARIN specifications the repository includes the possibility to upload CMDI metadata files. This is an administrator function performed by Språkbanken staff.

Like mentioned the repository software is the LINDAT/CLARIN modified DSapce available here:

https://github.com/ufal/lindat-dspace

The LINDAT/CLARIN repository has obtained the Data Seal of Approval.

https://lindat.mff.cuni.cz/en/

For Guideline 7 and the work on future-proofing obsolescence of file formats we collaborate with another University of Gothenburg entity, Swedish National Data Service (SND), with which we also share some staff, so this is not really an outsourced part.

SND has obtained the Data Seal of Approval.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

1. The data producer deposits the data in a data repository with sufficient information for others to assess the quality of the data, and compliance with disciplinary and ethical norms.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

We consider this as implemented.

The depositor is required to comply with the following:

- Data have to be provided with metadata in standards formats accepted /adopted in the respective communities.
- Persistent Identifiers (PIDs) have to be assigned.
- IPR issues have to be resolved and clear statements regarding licensing and possible use of the resources are to be made.
- The depositor is also required electronically sign a deposition agreement acknowledging being holder of rights to the data and has the right to grant rights contained in the license.

FAQ:

https://repo.spraakbanken.gu.se/xmlui/page/faq

"Data in the Språkbanken CLARIN Repository are made available under the licence attached to the resources. In case there is no licence, data is made freely available for access, printing and download for the purposes of non-commercial research or private study. Users ***must*** acknowledge in any publication, the Deposited Work using a 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 licence without formal permission of the copyright holders."

Also the user agrees "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."

Terms of Service:

https://repo.spraakbanken.gu.se/xmlui/page/about#terms-of-service

https://repo.spraakbanken.gu.se/xmlui/page/terms-of-service

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

2. The data producer provides the data in formats recommended by the data repository.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The data producer is encouraged throughout the submission process to use one of the recommended formats mentioned in formats for language resources and tools (LRT):

http://www.clarin.eu/sites/default/files/Standards%20for%20LRT-v6.pdf

this is linked from our FAQ where also examples of accepted resources are given:

https://repo.spraakbanken.gu.se/xmlui/page/faq#what-submissions-do-you-accept

Usage of any other data formats than the recommended in a submission will be followed by an interaction between the reviewer/editor and the data producer with advice for conversion before it can be accepted. If the data producer provides extensive enough documentation allowing data converters to be built the resource can also be archived in its original data format to avoid conversion loss.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

3. The data producer provides the data together with the metadata requested by the data repository.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The submission work flow consists of several steps where the data
producer must enter mandatory metadata as part of the deposit procedure:

https://repo.spraakbanken.gu.se/xmlui/page/deposit

During the submission we require that the user provides at least the following information:

- type of the resource - currently allowing only 4 types (corpora, tools, language conceptual resources, language descriptions)
- title
- list of authors
- issue date

- description

- publisher

- resource language(s) code(s) (if applicable)

- contact person (the responsible person for the submission information) - at least surname and email

- distribution information - access rights, licence information, licence restrictions, distribution media

- content information - type of media (eg. text/audio/...), (if applicable) further classification of the resource (e.g. ontology/thesaurus for lexical conceptual resources)

- size information - size in bytes/words/n-grams/... (if applicable)

Metadata (data) can be imported from other repositories which support standard protocols for sharing (e.g., OAI-PMH, OAI-ORE, DSpace Archival Information Package).

We employ a complex set of automatic curation tools which report the quality of metadata regularly to the Språkbanken repository administrator.

The automatic curation tools include validation, checks for selected values from semi-closed sets, checking for missing values and invalid values. Updates are fetched from authority databases to validate vocabulary values. Are author's, publisher's and data provider's names fully specified? Are email addresses valid? Is a language set for resources requiring one etcetera.

In case of missing or invalid data, submissions are immediately removed and the data provider is asked for improving the quality of
metadata before republishing the submission item.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The Språkbanken CLARIN Repository shares the mission statement of CLARIN ERIC:

"The ultimate objective of CLARIN ERIC is to advance research in humanities and social sciences by giving researchers unified single sign-on access to a platform which integrates language-based resources and advanced tools at a European level. This shall be 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 (henceforth abbreviated HSS) research communities at large."

https://repo.spraakbanken.gu.se/xmlui/page/about#mission-statement

The Språkbanken CLARIN Repository is a dedicated part of the Swedish Swe-Clarin and international CLARIN infrastructures. It is hosted and maintained at Språkbanken, Department of Swedish Language at University of Gothenburg.

The Swedish governmental digital assets preservation and access policies are applicable to all actions undertaken by Språkbanken as an organizational unit of University of Gothenburg a state owned legal entity. No part of the preservation is actually outsourced, but we do collaborate with Swedish National Data Service (SND) and also share some staff. SND has obtained the Data Seal of Approval.

Since we keep an automatic inventory of all data formats deposited and only allow the recommended formats specified in Guideline 2 and handle future-proofing of file format obsolescence in Guideline 7 the preservation should cover the full range of objects deposited.

We plan to officially launch our repository at a national opening ceremony October 7th here at Språkbanken in Gothenburg.

We promote the repository in the Swe-Clarin network at partner activities, open and directed, and at Nordic and European level CLARIN activities.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

5. The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The repository is an official part of Språkbanken, Department of Swedish Language, University of Gothenburg which is a legal entity owned by the state.

The Språkbanken CLARIN Repository requires submitters to electronically sign the right to archive the data and the that the responsibility of the content lies with them. This means the depositors are solely responsible for taking care of IPR issues before publishing data or tools by submitting them to the repository.

https://repo.spraakbanken.gu.se/xmlui/page/about#about-ipr

The repository states the following privacy policy:

https://repo.spraakbanken.gu.se/xmlui/page/privacypolicy

The repository complies with the Data Protection Code of Conduct:

http://geant3plus.archive.geant.net/uri/dataprotection-code-of-conduct/v1/Pages/default.aspx

Everyone who downloads data is bound by the licence assigned to the item – in order to download protected data, one has to be authenticated and needs to electronically sign the licence.

The repository requires data users to comply with the data citing policy:

https://repo.spraakbanken.gu.se/xmlui/page/cite

Also from Guideline 1 about terms-of-service:

https://repo.spraakbanken.gu.se/xmlui/page/terms-of-service

The user also agrees "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."

Data with disclosure risk is handled appropriately. The depositor is asked about sensitive information and licensing issues concerning the data. All data are reviewed by Språkbanken staff for disclosure risks and, when necessary, modified in consultation e.g. anonymization.

Knowledge and compliance with national and international law, such as the Swedish Public Access to Information and Secrecy Act (Offentlighets- och sekretesslagen 2009:400) and the Personal Data Act (Personuppgiftslagen 1998:246) is ensured by internal training and counselling.

Data at Språkbanken are stored, managed and distributed according to the Regulation for IT-Security at University of Gothenburg:

http://medarbetarportalen.gu.se/digitalAssets/1531/1531415_regulations-for-it-security_-revision2015_rev_pl.pdf

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

6. The data repository applies documented processes and procedures for managing data storage.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

Data storage is done in the LINDAT/CLARIN modified DSpace software constituting the Språkbanken CLARIN Repository.

https://wiki.duraspace.org/display/DSPACE/User+FAQ#UserFAQ-HowdoesDSpacepreservedigitalmaterial?

As a part of a preservation strategy “DSpace allows you to identify two levels of digital preservation: bit preservation, and functional preservation”.

The Språkbanken CLARIN repository is backed up daily by Språkbanken. This includes the database and the resource folder as well as the repository source code and installation files. Disk storage is a high availability 3 node distributed storage cluster. Backups are preformed by Bacula backup system. All backups are checksummed for consistency.

We have a data retention policy of 3-6 months for direct on-line backup availability. Every third month snapshots are made in two copies which are then stored off-line in two different premises security sections, and always different fire safety sections. Every third year these copies are transferred onto and verified on new media.

Data at Språkbanken are stored, managed and distributed according to the Regulation for IT-Security at University of Gothenburg:

http://medarbetarportalen.gu.se/digitalAssets/1531/1531415_regulations-for-it-security_-revision2015_rev_pl.pdf

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

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

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The digital assets of the repository will be long-term preserved by Språkbanken for availability and we collaborate with Swedish National Data Service (SND) another University of Gothenburg entity for future-proofing obsolescence of file formats. We also share some staff with SND. No part of the preservation is actually outsourced.

The repository is based on DSpace repository system which is one of the leading software in this category. DSpace supports state-of-the-art preservation tools in various forms. From simple replication to standard backup formats and easily manageable collections. The metadata can be exported into many various formats suited for long time preservation including self describing ones like XML. Multilingual support is secured by using Unicode at every level. The XML format is used at several occasions e.g., when exporting to specific CMDI (Component MetaData Infrastructure) profile or when archiving AIP (Archival Information Packages). The format validation is done regularly using external harvesting service (http://validator.oaipmh.com/); moreover, there are several institutions which harvest our repository regularly and these make the validation too.

Språkbanken minimizes the cases in which obsolescence of file formats occur by using migration strategies. This includes only allowing usage of recommended formats specified in Guideline 2. Usage of any other format in a submission will be followed by an interaction between the editor and the data producer with advice for conversion before it can be accepted. If the data producer provides extensive enough documentation allowing data converters to be built the resource can also be archived in its original data format to avoid conversion loss.

An automatic inventory of the file formats present in our repository gives us a good overview of what file formats are used informing the migration strategies.

The item lifcyle is described here:

https://repo.spraakbanken.gu.se/xmlui/page/item-lifecycle

Development and progress in technologies are followed closely by active involvement in national and international standardization and interoperability work, e.g. SIS/TK 115:

http://www.sis.se/Templates/SIS/Pages/ProductTechnicalCommitteeView.aspx?id=37&epslanguage=en&pid=TC-78791&icslvl1=SIS_C

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

8. Archiving takes place according to explicit work flows across the data life cycle.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The repository is run by Språkbanken. Like LINDAT/CLARIN we distinguish between known submitters and unknown ones, where the submissions from the latter ones will be specially validated and verified. A known submitter is one authenticating via its home institution or one who has been verified with a local account. The submission work flow is internally configured in the repository and the submitter goes through each step of it.

The repository has a curation process and a procedural documentation for archiving data:

https://repo.spraakbanken.gu.se/xmlui/page/deposit

After the submission our editors get the submission.

We have automatic tools helping the editors to verify and validate metadata and the integrity of the submitted data which are performed by every editor during the curation step and automatically at regular time intervals.

The repository also documents life-cycle questions like editing/modifying or deleting data:

https://repo.spraakbanken.gu.se/xmlui/page/item-lifecycle

If submissions happen to not be part of the Språkbanken CLARIN/CLARIN ERIC mission advise on where to direct such submissions will be given.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

9. The data repository assumes responsibility from the data producers for access and availability of the digital objects.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The author/creator of the resource remains the proprietor. Språkbanken assumes responsibility for accessibility and availability of the repository in accordance with Swe-Clarin.

Data providers have to state a licence as part of the submission procedure otherwise it cannot be submitted. We guide data providers on request to use as few licences as possible and primarily to choose one of the available CC ones.

The repository is covered by the Språkbanken backup routines and crisis management (see also Guidelines 6-7). The risk of data being lost due to minor or major crises is very low. Virtual machines are used which can be migrated by the Språkbanken system administrators to other physical hardware.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

10. The data repository enables the users to discover and use the data and refer to them in a persistent way.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The repository is browsable and searchable. Since it is also harvested by other centres like the CLARIN Virtual Language Observatory (VLO) the metadata will be visible and discoverable in those too. References can be persistently made to PIDs provided by the repository. The repository also generates full bibliographical references on the fly.

The repository provides various ways of utilising the archived data via on-line tools as well as by downloading the data in formats commonly used by the research communities:

https://repo.spraakbanken.gu.se/xmlui/page/faq#what-submissions-do-you-accept

The repository is stable. Språkbanken runs its own Handle server providing PIDs.

OAI and other harvesting options are permissible and available.

There are advanced metadata search capabilities in the repository and one-click links to search the contents of discovered resources in CLARIN Federated Content Search.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

11. The data repository ensures the integrity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

MD5 checksums are utilised by the underlying DSpace software for all objects. The repository also monitors data and metadata integrity regularly. The availability of files, web and application servers is monitored continuously.

Multiple versions of data are handled. Once deposited, files in an item's data set can not be changed by the submitter but only by administrators.

The assigned persistent identifiers always refer to the same content. Currently, if a submission is superseded by a new version the old one is withdrawn. This means that it cannot be searched for and that it is not displayed in any statistics. However, the PID url is working showing the
submission as before with a special metadata value *isreplacedby* added which points to the new version. See Deleting and Modifying of Published Item in life-cycle:

https://repo.spraakbanken.gu.se/xmlui/page/item-lifecycle

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

12. The data repository ensures the authenticity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The repository's item life-cycle information makes data producers aware of the strategy for data changes:

https://repo.spraakbanken.gu.se/xmlui/page/item-lifecycle

Provenance metadata are stored in log messages for every change.

Data providers cannot change a submitted item's files or metadata without contacting the repository administrators. Minor edits of metadata, e.g. fixing typos, can be allowed. For non-trivial changes a new version is required and indicating in metadata that the old item is replaced. Removals always keep the metadata and the PID resolves to it if only to inform about the item being deleted. The importance and feasibility is being evaluated per case.

Depositors are only users that are authorised by their home institution's federated log-on provider or identified and verified local users.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

13. The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

We are committed and rely on the group of emerging standards around CMDI component metadata model (ISO-CD 24622-1) for metadata standards:

http://www.iso.org/iso/catalogue_detail.htm?csnumber=37336

Based on the LINDAT/CLARIN modified DSpace software and the defined work flow supported by the repository's interface, the Språkbanken repository meets the requirements of OAIS as described below.

http://registry.duraspace.org/about

1) Ingestion: The Submission Information Packages (SIPs) are received for curating and are assigned to a task pool where our curators can process them. There is a number of pre-configured supported SIP formats

https://wiki.duraspace.org/display/DSDOC5x/Importing+and+Exporting+Content+via+Packages

However, the default way is that the ingestion process is done through our web based interface which hides the implementation details.

2) Archival Storage: Using the web interface a curator updates (adds, deletes, modifies) the metadata and the submitted bitstreams are validated. In general, the curators ensure consistency and quality of each submission. If the curator approves an item, the Archival Information Packages (AIPs) is available.

3) Data Management: This function is executed during the creation of the metadata (descriptive, administrative and structural), as seen on the prior step.

4) Preservation Planning: As described in section 6, we monitor and backup our system in several layers. More preservation details are described in section 9. In a repository context, each submission bitstream has md5 checksums which are regularly checked. 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).

5) Administration: In general, there is no administration of Data Producer prior to submitting and item. We are open to all submissions which meet our standards (Data Producers must be authenticated which means they must have academic background or have verified local accounts). A contract is signed during the ingestion process. A specific robust administration interface is available including specific detailed reports on the contents of our repository.

6) Access: The available Dissemination Information Package types

https://wiki.duraspace.org/display/DSDOC5x/Importing+and+Exporting+Content+via+Packages

query responses and reports are delivered to CONSUMERS. Few submissions require authenticated access which is granted to academic users (through shibboleth) and locally registered users. Few submissions have their bitstreams available after specified date. DSpace allows for searching, locating and description of the information stored. All metadata are publicly available.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

14. The data consumer complies with access regulations set by the data repository.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

Access regulations are based on federated log-on systems, by which consumers are authenticated. We offer local accounts too and in this case we perform the identification and verification manually. In case of stricter licences for certain data sets, a contract is provided in the form of a click-through acceptance of licences. Acceptance of licences is logged by the repository and available for administrator inspection.

Each submission is clearly marked with its licence. The metadata themselves are always public. In any case, we strongly encourage data publishers to use CC licences or at least one already available licence.

Usage of downloaded data is not monitored. Only download time and date, the identity of the consumer and the documented acceptance of the licence.

https://repo.spraakbanken.gu.se/xmlui/page/item-lifecycle

If a report on suspected IPR infringement comes in the documented acceptance of the licence by the consumer can be released to the data provider or proper authority on request.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

15. The data consumer conforms to and agrees with any codes of conduct that are generally accepted in the relevant sector for the exchange and proper use of knowledge and information.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The data consumer pledges to conform to and agree with the general codes of conduct that are accepted by the data consumer when she is granted access within a national federated log-on system for the academic sector.

Data providers are required explicitly to ensure that IPR and personal rights are respected in their data.

The licence of each item is clearly stated.

The ethical terms of service are also clearly stated by the repository:

https://repo.spraakbanken.gu.se/xmlui/page/terms-of-service

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

16. The data consumer respects the applicable licences of the data repository regarding the use of the data.

Minimum Required Statement of Compliance:

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

Applicant Entry

Statement of Compliance:

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

Self-assessment statement:

The data consumer is expected to respect the general codes of conduct. The data consumer is further expected to explicitly acknowledge any stricter licence that might be applicable to certain data.

The data consumer is made aware of usage restrictions using clear visual indicators, e.g.,

This item is

Publicly Available

and licensed under:
Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0)

If the data are licensed with a licence that requires signing, the user is asked to electronically sign the licence before downloading. The Språkbanken CLARIN Repository currently only have items deposited with publicly available licences. Examples of licences requiring signing would be "academic use" or "restricted use".

In case of misuse, the only thing that can be practically done is to deny the user further access to the repository and to make the research community aware of the misuse. Each signing by the user is stored as well as the time and

date of any download of the resources.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments: