



## **Assessment Information**

[CoreTrustSeal Requirements 2017–2019](#)

Repository: GAMS - Geisteswissenschaftliches Asset Management System  
Website: <http://gams.uni-graz.at>  
Certification Date: 18 April 2019

This repository is owned by: **Zentrum für Informationsmodellierung**



# GAMS - Geisteswissenschaftliches Asset Management System

## Notes Before Completing the Application

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

True

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

.

## CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

### Background & General Guidance

### Glossary of Terms

## BACKGROUND INFORMATION

### Context

*R0. Please provide context for your repository.*

## ***Repository Type. Select all relevant types from:***

Domain or subject-based repository, Institutional repository, Library/Museum/Archives, Research project repository

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## ***Comments***

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

#### **Reviewer 2**

Comments:

## ***Brief Description of the Repository's Designated Community.***

GAMS is an asset management system for the management, publication and long-term archiving of digital resources from the Humanities. It enables scholars, researchers and students to manage and publish project resources with permanent identification and enriched with metadata.

Design and development of GAMS are carried out by the Centre for Information Modelling at the University of Graz in cooperation with multiple partners inside and outside the university, with regards to the specific requirements of humanistic research.

The repository is fully OAIS (Open Archival Information System)-compliant and covers the full life cycle of digital objects from receiving the SIP (submission information package), archiving the AIP (archival information package) and delivering the DIP (dissemination information package) to the public.

The designated community consists of scholars and students of the Arts and Humanities and related disciplines, as well as cultural heritage institutions (libraries, archives, museums).

Data can only be deposited in the course of a research project collaboration with the Centre. Each project is assigned a metadata manager from the Centre to assist with the workflow, data modelling, deposition and publication processes of the project. This approach guarantees the creation of high quality data prepared for scientific publication and long-term archiving. This also means that data is always at least basically curated when ingested into the repository. Project partners can come from the Faculty of Humanities or the University of Graz, but also from other local, national or international research or cultural heritage institutions.

Cf. also the collection policy and description [here](#)

<http://gams.uni-graz.at/archive/objects/context:gams/methods/sdef:Context/get?mode=about&locale=en>.

*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, D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

*Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

***Comments***

*Reviewer Entry*

**Reviewer 1**

Comments:

**Reviewer 2**

Comments:

***Outsource Partners. If applicable, please list them.***

*Reviewer Entry*

**Reviewer 1**

Comments:

**Reviewer 2**

Comments:

## *Other Relevant Information.*

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

#### **Reviewer 2**

Comments:

## **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:*

In the published mission statement of the Centre (German,

[https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/2016\\_Gruendungserklaerung\\_ZIM\\_V2.pdf](https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/2016_Gruendungserklaerung_ZIM_V2.pdf)

) long-term archiving and preservation of research data in a digital repository are stated explicitly as one of the main research themes of the Centre (first point in §2, p. 1).

Furthermore, „Betrieb eines Digitalen Repositoriums (Geisteswissenschaftliches Asset Management System GAMS) und einer forschungsbezogenen IT-Infrastruktur zur Langzeitarchivierung von (geisteswissenschaftlichen) Forschungsdaten.“

(third point in § 8, p. 4) Translation: “operation of a digital repository (GAMS) and a research-related IT infrastructure for long-term preservation of (Humanities’) research data“ is listed as a main task.

This is acknowledged by the University, and the Centre’s infrastructure partly acts as an institutional repository for the

Faculty of Arts and Humanities in Graz (cf. the wide variety of projects and collections at <http://gams.uni-graz.at/>).

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept.

Future submissions could possibly offer the primary mission docs in English.

## **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:***

Data can only be deposited in the course of a cooperation project. Therefore, issues like IPR (Intellectual Property Rights) and licensing (possible limitation of access) are discussed and determined at the very beginning of project planning. The Centre also offers consulting and expertise in these fields. Project partners are responsible for respecting respective national and international laws. Depending on the project, suitable agreements on these legal terms are signed, but – due to the diversity of the data and the project partners – not in a standardized way.

As indicated in its mission statement, the Centre promotes open access and free availability of research data. All resources (facsimiles and metadata) of the GAMS repository are licensed under a Creative Commons license, usually CC BY-NC (<https://creativecommons.org/licenses/by-nc/3.0/at/>) or CC BY-NC-SA (<http://creativecommons.org/licenses/by-nc-sa/4.0>). This is indicated on the repository website (cf. <http://gams.uni-graz.at>) and clearly displayed also at the individual object level. Usually, a global project license in the form of a graphic is featured

in the HTML footer and an individual license in the form of a link is featured in the object metadata (cf. for instance this example: <https://gams.uni-graz.at/o:km.9>). Both links refer the user to the respective Creative Commons license in English, thus ensuring that non-German speakers can also access license information.

Upon request of the cooperation partners who act as data providers, access regulations to their digital objects can be put in effect. Fedora supports management of access right via XACML (eXtensible Access Control Markup Language). Another possibility is to use a password for the web representation, which can be requested by the users directly from the data providers.

In case of non-compliance with the regulations determined by the data providers (for instance commercial use of non-commercially licensed material), the publishers are free to take legal action against the offenders. The repository will bring such cases to the attention of the providers if they gain knowledge of it and support the actions decided by the data providers.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

### **III. Continuity of access**

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

#### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

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

##### **Reviewer 2**

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

#### ***Response:***

The University of Graz recognizes its responsibility to preserve and maintain the data of the GAMS repository now and in the future. The permanent and continued access to over 50 data collections from the fields of humanities and cultural heritage is of central concern not only to the Centre for Information Modelling but considered an obligation by the whole University. To further emphasize this concern, the organizational unit responsible for this task (ZIM-ACDH) is currently in the process of being transformed from a temporary Centre to a permanent department (Institute) of the University of Graz. The operation of the GAMS repository will remain to be one of the main duties of this Institute, in addition to continuing research in the Digital Humanities.

This commitment is also further underlined by the fact that since 2012 the number of employees has doubled and the university has set up a professorship for Digital Humanities in 2016. Austria's first MA degree in Digital Humanities was implemented in 2017.

If the University of Graz should in the future for whatever reason not be able to fulfill this duty, it will ensure the continuity and sustainability of the data concerned. A data package containing the data in its entirety and in a self-descriptive archival format will be exported from the repository and either be returned to the original owner or transferred to another suitable disciplinary or institutional national repository.

Usually, GAMS is the only custodian of the research data in question and recognizes the importance of securing the preserved data, be it in their own custody or by guaranteeing the transfer to a different appropriate environment or infrastructure. To further facilitate the transfer to other repositories if the worst comes to the worst, there are continuous efforts to create a dedicated Austrian repository federation, where existing repositories (especially the ones operating on the same base layer Fedora Commons, like ARCHE (<https://arche.acdh.oew.ac.at/browser/>) or PHAIDRA (<https://phaidra.univie.ac.at/>)) could act as a fallback solution to each other. This has been repeatedly discussed at various occasions, for instance in the working group "archival" of the Austrian federation project KONDE (<http://www.digitale-edition.at/>) for which the Centre acts as project lead.

[Since the Centre is currently in the process of being transformed from a temporary Centre to a permanent department (Institute) of the University of Graz, the Vice Rector for Research has underpinned the commitment of the University to the above statement in a letter attached to the application.]

#### *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:***

Depositors are responsible for adherence to national and international law as well as respect of ethical norms and discipline specific rules. This responsibility is also specified in the deposition agreement. If personal or sensitive data with a disclosure risk are part of the project, there will be provisions in place to eliminate or minimize this risk. One possibility is to pre-process the ingested data accordingly; another is that the project-specific ingest workflow will consider the possibility of confidential data and employ suitable measures. This includes for instance anonymization, e.g. blurring of names and addresses on postcard facsimiles (<https://gams.uni-graz.at/o:gm.100> ) or blurring of names on historical criminal records (<http://gams.uni-graz.at/o:km.9>) for the facsimiles. With regard to textual (meta)data, the data creation workflow (usually consisting of XSL transformations) can specifically query for confidential data (like personal names) and eliminate or anonymize them. These measures should prevent the ingest of confidential data altogether, also due to the fact that ingest and data modelling are carried out by a metadata manager individually assigned to the project. So far, no written procedures have been needed to address this issue, and confidential data has not been included in the repository by accident. If such an event should occur in the future, the repository managers would be prepared to immediately exclude the respective data.

Usually, IPR is less of an issue in our case, because the data we are dealing with is usually beyond the protection time limits; if necessary, only restricted access to the material will be granted (cf. also the section on Licenses/R2).

### ***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:***

For the current period of existence the Centre is sufficiently equipped with funding and qualified staff. Currently, five staff members with an equivalent of 3,5 full positions are permanently employed to take care of various aspects of the infrastructure, covering not only technical but also organisational and content-related tasks. Occasionally, this team is supplemented by short-term funded staff responsible for individual additions or dissemination methods. All staff members have the opportunity to attend suitable training courses and relevant conferences (see also R6).

In general, requirements for the staff usually include a combination of Humanities and IT skills, i.e. a BA or MA degree in a Humanities discipline alongside advanced IT skills (formalized training on secondary or tertiary level, work experience, supplementary courses). A substantial portion of staff is recruited from DH teaching at the University of Graz, formerly the Joint Master's Degree EuroMACHS (<https://euromachs.uni-graz.at>), now the MA Digital Humanities (<https://informationsmodellierung.uni-graz.at/en/studying/master-degree-digital-humanities>).

However, the high volume of project-based funding acquired by the Centre results in an asymmetry between temporarily and permanently employed staff members of the repository, which calls for adjustments in the training and number of permanently employed staff members. This has to be a continuously addressed process and can never be regarded as “fully implemented” or “finished”. It is one of the main risks of maintaining a trusted digital repository, which can never be fully eliminated and is of particular consequence when a high percentage of the permanently funded working hours is already bound by large and complicated tasks such as the migration of the storage layer. The sudden loss of key staff members is also an inevitable risk, which we seek to reduce by providing and curating sufficient documentation, as well as knowledge transfer among the employees (see also R16 on risk management).

#### ***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:***

Scientific guidance is partly obtained from the scientific advisory board of the Centre (<https://informationsmodellierung.uni-graz.at/en/centre/scientific-advisory-board/>). The board consists of two national and three international members, seeking to represent also the two main research perspectives of the repository, namely Digital Edition (currently Claudine Moulin) and Digitized Cultural Heritage (currently Lorna Hughes). All members contribute to an enhanced scientific network on a national and international level, being well represented in their respective communities and associated committees (for instance dha, DHd, etc.).

A key element of guidance and sustainable networking is the knowledge transfer within the Centre itself. Various staff members have affiliations with relevant organisations and communities (TEI Council and Board, Research Data Alliance, DARIAH). Internal communication is facilitated by a designated “research meeting” every Wednesday, where current projects and developments are presented and discussed. Another possibility for networking and exchange with colleagues from other institutions or disciplines is the attendance of designated working groups (Working Group on data centres of the DHd, Research Data Alliance, CLARIN and DARIAH, Fedora User Group, etc.) and special conferences. The Centre also seeks to complement their skills by inviting experts to our “lunchtime lectures” (<https://informationsmodellierung.uni-graz.at/de/veranstaltungen/lunchtime-lectures/>) or organizing international conferences, workshops and summer schools (<https://informationsmodellierung.uni-graz.at/de/veranstaltungen/>).

*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:***

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:***

Fedora Commons supports versioning of every aspect of the digital resource from primary source to metadata and associated materials. All changes and previous versions of the material can be retrieved and are accessible from the infrastructure.

The Cirilo client checks well-formedness of XML-formats and validates against the referenced schema upon ingest. This applies also to metadata, as it is stored as an XML-based datastream within the digital object.

Fedora uses MD5 checksums to guarantee the integrity of the resources in the digital archive. This operation is carried out each time new material is ingested or a resource is modified, as well as regularly once a month to compare all datastreams and generating a report. For PDF and image formats FITS (File Information Tool Set, <https://projects.iq.harvard.edu/fits>) is in use.

Due to the special nature of cooperation projects, a large portion of personal contact is needed with the depositors and their identity can be established from the beginning.

Although GAMS accepts responsibility for providing data integrity and authenticity and is currently already guaranteeing to do so, the compliance level is set to “in implementation phase”. This is because of the sometimes lacking transparency of these processes to the end user, e.g. previous versions of the data can basically only be accessed with knowledge of the infrastructure and mostly not transparently offered on the surface of the project to the audience. This is something we are aware of and seek to improve especially with regard to the migration to Fedora 4 currently underway. The goal would be to offer the so far largely hidden information transparently to the audience by a standardized module, thereby also improving the scientific quality and transparent and replicable research process reflected in the data.

#### *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:***

Data can only be deposited as part of a cooperation research project. Such a project usually entails a joint funding application at national or international funding agencies. Thus, the Centre for Information Modelling and the project partner jointly define a research question and establish the (digital) research methods used in the project. The application is supplemented by a data management plan detailing all information on the expected research data to be integrated in the repository. In case of approval, the proposal and integrated data management plan form the basis of the formal cooperation and deposition agreement signed between the partners. The research question often comes from the primary research objectives of the Centre, namely Digital Edition and Digitized Cultural Heritage, but other proposals are also welcome to broaden the repository's range of data formats and support the integration of new functionalities.

Each project is supported by a metadata manager from the Centre who assist the project partner with the workflow, data modelling, deposition and publication process in the course of the project. This approach guarantees the creation of high quality data prepared for publication and long-term archiving. Metadata is provided in sufficient detail for both purposes. Quality control is especially supported by the use of controlled vocabularies and authority files. Every digital object must by definition be equipped by a Dublin Core datastream containing sufficient data for information retrieval and resource identification.

GAMS is serialized on XML (eXtensible Markup Language) -based standards and technologies for data storage and representation. If the data in question does not conform to any XML-based international standard, the Centre will implement suitable workflows for the conversion of the content in agreement with the project partners.

For text and metadata, the Centre uses (among others) the following standards: TEI (Text Encoding Initiative), LIDO (Lightweight Information Describing Objects), DC (Dublin Core), METS/MODS (Metadata Encoding and Transmission Standard/Metadata Object Description Scheme), RDF (Resource Description Framework), SKOS (Simple Knowledge Organization System). This list of preferred formats is reflected in the use of dedicated content models for the respective standards (<http://gams.uni-graz.at/o:gams.doku#cirilomodels>). The Cirilo Client then checks the well-formedness of the XML and validates the document against the given schema to ensure conformity. Data producers must deliver images in the recognized standards JPEG/JPEG2000 (Joint Photographic Experts Group) or TIFF (Tagged Image File Format); if necessary the Centre assists in a conversion process.

In our opinion, operating with a limited set of preferred formats and data types is key to keeping a repository maintainable over a long period of time, thus, compliance with standards and preferred formats is closely monitored.

#### *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:***

The archival storage of the repository has sufficient measures for storage, identification, retrieval and backup in place. These measures pertain to all levels of security, be it a) data security in the sense of authenticity and integrity (see R7), b) authentication for designated users only (for depositors and managers as well as for users) and c) network and system security itself (see R16).

Data storage is provided via SAN by our university's IT department (UNI IT). Data is stored redundantly in two data centers in different campus buildings. A formal service level agreement between the Centre for Information Modelling and the IT department as hardware and data provider is currently in preparation.

Data backup in GAMS is part of the central backup processes of the University. Backups run daily and are stored on a disk array and later moved to tape. There is an additional offsite backup managed by the Centre which is also run every night. The combination of both strategies ensures that backups are available for seven years. Data recovery is regularly exercised on a spare machine for training purposes of the administrators.

Backup consistency is guaranteed because every Fedora object is entirely stored in FOXML format containing all binary data streams in base64 encoding. Additionally, all datastreams are preserved in the original format as distinct files. These processes are documented in detail in the Fedora Commons wiki (<https://wiki.duraspace.org/display/FEDORA/All+Documentation> and select your version).

Only system administrators have direct access to these data, all other operations performed on the data are performed via the Cirilo client requiring authentication via LDAP. As each object provides MD5 checksums for the datastreams, corrupted data can be identified easily.

All services are monitored 24/7 with a reaction time of typically only a few minutes during working hours to some hours during weekends and nights.

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept.

For the next renewal, it would be good to see that the formal service level agreement has been executed between the

Centre for Information Modeling and the IT department.

#### **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:***

So far, the repository does not have a dedicated preservation plan in the form of a separate document: rather, the relevant contents are distributed through various sources. Responsibilities and obligations of the contractors are listed in detail in the cooperation and deposition agreement signed with the project partners at the beginning of the cooperation (sample available here: [https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/Repository-Depositors-Agreement\\_GAMS\\_V3.pdf](https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/Repository-Depositors-Agreement_GAMS_V3.pdf)). Applying for the related project funding usually also means handing in a designated data management plan (sample available here: [https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/Data\\_Management\\_Plan\\_GAMS-Sample.pdf](https://static.uni-graz.at/fileadmin/gewi-zentren/Informationsmodellierung/PDF/Data_Management_Plan_GAMS-Sample.pdf)). More general principles are outlined in the About section (<http://gams.uni-graz.at/archive/objects/context:gams/methods/sdef:Context/get?mode=about&locale=en>) and the documentation of the infrastructure (<http://gams.uni-graz.at/docs>). This especially pertains to the list of preferred formats suitable for preservation, the Fedora object model structuring the data and the data life cycle covering the all functions of the OAIS-model. By definition, the underlying software Fedora Commons is an OAIS-compliant asset management system, addressing and managing all stages of the data life cycle.

The Centre assumes full responsibility for the long-term preservation of the repository content, as stated in its mission statement (see R1). All components of the infrastructure are continuously monitored as part of preservation planning. Procedures are put in place to further guarantee long-term preservation, most notably in performing necessary migration processes. It is our belief that only by making use of a restricted number of formats and technologies and applying



standardized workflows a repository can be maintained over a long period of time and take on the responsibility of long-term preservation. This principle forms the basis of all decisions in preservation planning.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept.

For the next renewal, it would be good to see a preservation plan.

##### **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:***

Since data can only be deposited in a cooperative effort, quality control in the curation of data and metadata is an integral part of a project. The metadata manager will not only choose the appropriate standard to represent the deposited information (e.g. TEI for texts, LIDO for museum objects, etc.) and make sure that the documents are well-formed and valid (see R7 and R8 for details), but will also check the data quality with regard to the content. This will include integration of suitable vocabularies and norm data on the formal side, but especially comprise constant communication and training of the scholars responsible for data creation and enrichment. This point has turned out to be crucial for data quality and is therefore fostered by regular project meetings and informal communication among the employees,

whenever possible. The metadata managers will usually have a humanities degree and be specifically trained to fill their role in bridging the gap between information science and traditional humanities scholarship. The data and metadata created in this way is directly accessible to the end user, not only through a graphical user interface but usually also by linking to the primary data. This includes information on the publisher responsible for the scientific content and often also the scholar responsible for the data creation. In many cases, there is a direct link to contact the editors and/or publishers with questions, remarks or corrections.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept.

For the next renewal, it would be good to see that citations to related works or links to citation indices are provided.

##### **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:***

An overview of the data life cycle in the repository can be viewed in the figure and description here:

<http://gams.uni-graz.at/archive/objects/context:gams/methods/sdef:Context/get?mode=about&locale=en>.

Archiving takes place following specific procedures laid out in R9. In cooperation with the project partners, the Centre takes over the data according to the specifications of the repository, with special emphasis on the list of accepted formats and data quality. This SIP is then ingested using the Cirilo Client. During ingest, metadata is extracted from the source

and mapped to a DC record. Additionally, semantic enrichment (like resolution of place names and ontology concepts based on authority files or thesauri) can take place. Images and other related materials are bundled within the resource automatically. This also includes validation and quality control of the data as well as assignment of a PID (persistent identifier) in the system.

The GAMS repository is not only used for long-term preservation but also for the web-representation of the resources. In that respect, it takes advantage of Fedora's object model, which assumes that all digital assets are completely self-descriptive (cf. <https://wiki.duraspace.org/display/FEDORA36/Fedora+Digital+Object+Model>).

The production of DIPs (for instance via stylesheets and transformations) is directly encapsulated in every resource. Thus, the GAMS repository and the Cirilo Client cover the whole lifecycle of objects from SIP to DIP in the OAIS model. These workflows are based on the inherent features of Fedora Commons and its object model, which are both outlined in detail in the documentation (cf. <http://gams.uni-graz.at/docs>).

Since the Centre is involved in any project from the application stage, the suitability of and compatibility with data for the repository is clearly established from the beginning. If the data in question does not fit the focus of the repository (i.e. arts, humanities and cultural heritage) or is deemed unsuitable for other reasons (for instance legal issues), a deposition and cooperation agreement can possibly not be reached. In this case we may assist the depositor in finding a more suitable solution for their data, for instance by establishing contact with a different institution.

#### *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:***

For users, the first point of access will usually be the graphical user interface created specifically for the project in question. It offers suitable ways of interacting with the respective data. There is also the possibility to perform a full-text search on the complete data basis of GAMS with the search field at <http://gams.uni-graz.at> and to make use of Dublin Core related queries. In addition, you can browse the Fedora backend at <http://gams.uni-graz.at/archive/search>. To systematically harvest or integrate data, the OAI-PMH interface at <http://gams.uni-graz.at/oaiprovider> can be used and queried according to the PMH. Currently, three metadata prefixes are supported: oai\_dc (Dublin Core), oai\_europeana (Europeana Semantic Elements) and oai\_edm (Europeana Data Model).

With regard to images, all image data stored in the repository can be delivered via a IIIF-compatible image server. To this end, use the syntax <http://gams.uni-graz.at/iiif/{object PID}/{image PID}/{IIIF query}>.

In agreement with the partners, data can also be contributed to various aggregation services and metadata registries like Europeana, Kulturpool, Pelagios or Nomisma, supporting discipline-specific ways of data discovery.

All resources in the repository have a PID and are addressable with the permalink <http://gams.uni-graz.at/PID>.

Datastreams can be accessed in the same way with <http://gams.uni-graz.at/PID/DATASTREAM>. This assures direct access, quotability and persistent identification for scientific contexts (cf. <http://gams.uni-graz.at/doku#pid>).

The Centre is member of the Handle network and is running its own handle server. Our handle prefix is 11471. This persistent identifier is stored as part of the objects metadata and of course published in the handle infrastructure (cf. <http://gams.uni-graz.at/o:gams.doku#handles>).

A recommendation for data citation is given explicitly in some of the projects, but not always included in the presentation interface; the necessary information is nevertheless always represented.

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:  
Accept.

For subsequent renewals, it would be good to standardize the provision of recommended citations for each data product.

#### **Reviewer 2**

Comments:  
Accept

## **XIV. Data reuse**

***R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.***

## ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

#### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

## ***Response:***

The XML-based content strategy ensures the re-usability of data and metadata for humans and machines. The reduced list of accepted formats and standards (cf. R8) is considered suitable for long-term preservation and archiving purposes. A data object is only complete with the obligatory Dublin Core dataset attached to it, as well as information on the possibilities for re-use (see the licensing guidelines in R2) and technical and administrative metadata. A crucial point for re-use are sufficient possibilities for data discovery, especially with regard to the designated communities; thus we seek to disseminate the metadata through appropriate registries listed in R13. Information on the data creation usually forms a part of the resource, e.g. the description of encoding or annotation rules in the TEI header. Data and metadata will be migrated, if the need arises in preservation planning; for instance a conversion from older Europeana Semantic Elements format to the newer Europeana Data Model format.

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **TECHNOLOGY**

### **XV. Technical infrastructure**

***R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated***

## *Community.*

### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### *Response:*

The repository runs on virtual servers in the university's VMware cluster and is operated by an up-to-date version of Debian Linux. Data is stored in a SAN provided by the IT department (UNI IT) of the Graz University. Both the VMware cluster and the SAN have a redundant setup located in two data centers in different campus buildings.

The repository is completely built on open source, platform independent and community driven software projects. GAMS was conceived and developed on the basis of the OAIS-compliant open-source project Fedora Commons (Flexible Extensible Digital Object Repository Architecture) (<http://fedora-commons.org>) and has been continuously improved in the course of cooperative projects, addressing the specific needs of research.

Further integral technologies include Apache Tomcat, Apache HTTP Server, Apache Cocoon, Blazegraph graph database, Apache Lucene, Apache Solr, PostgreSQL database server, Loris IIIF image server and a Handle server.

The bundled infrastructure is available as an archive-in-a-box solution from GitHub (<https://github.com/acdh/cirilo>).

For dissemination purposes, the repository relies on HTML5, CSS, and JavaScript, all recommended W3C standards and validated upon publication of a project.

With regard to connectivity, the repository has a redundant 10-GigaBit connection via ACONet til Vienna.

Infrastructure development is largely guided by demands from preservation planning, i.e. obsolescence of software solutions and formats as well as continuous upgrade to newest software versions. Currently, this results for instance in a major migration project from Fedora 3.x to Fedora 4.

#### *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:***

Backup, network security and physical security are part of the university's security strategy handled by the university's IT service (Uni IT). Data is stored redundantly in two data centers in different campus buildings. Backups run daily and are stored on a disk array and later moved to tape. There is an additional offsite backup managed by the Centre which is also run every night. The combination of both backups ensures their accessibility over a period of seven years (see R9). Software security and authorisation management via LDAP is carried out by the system administrators and security officers of the repository. In case of failures or problems of one or more components of the infrastructure, the administrators are also notified via email to provide a reaction as swift as possible. Nevertheless, technical infrastructure is only as good and reliable as the staff maintaining it, thus loss of key staff members and sufficient funding for system administration is recognized as a main risk. Counter strategies like internal documentation, knowledge transfer and redundant coverage of tasks by different staff members (cf. also R5) seek to minimize this risk. Formalized risk analysis tools are currently not in use.

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

Accept.

For the next renewal, it would be good to develop a business continuity plan.

##### **Reviewer 2**

Comments:

Accept

# APPLICANT FEEDBACK

## Comments/feedback

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

### *Response:*

a lot of overlap between the questions

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Thank you for your feedback.

#### **Reviewer 2**

Comments: