Data Management Plan for GAMS Projects

Administrative Data

Project title	…
Acronym	…
Principal Investigator / Researcher	…
ID of Principal Investigator	ORCID:
Project Contact	… Phone: … Mail: …
DMP Contact	Zentrum für Informationsmodellierung – Austrian Centre for Digital Humanities (ZIM-ACDH) Phone: +43 (0)316 380 - 2292 Mail: zim@uni-graz.at
Funding Authority	…
Funding program	…
Grant Reference Number	…
Relevant policies	…
Date of initial DMP version	05.03.2018
Date of current DMP version	v.1: 05.03.2018

Project description

Provide a short description of the project

Data Collection

Primary data (text, images, audio, video)	*Describe digitization / generation of primary material*
Meta data	*Describe generation of metadata*
Analysis data	*Describe generation of analysis data*

Description of research data

Text	
Office documents	*Describe the handling of Office Documents.*
Hierarchical text data	Primary texts, comment and analysis data are stored in TEI P5 XML data streams (*.xml). *Describe the handling of other hierarchical text data.*
Program code	Data transformations and program code for presentation on the Web are created using XSL Transformation (* .xsl). *Describe the handling of other programme code.*
Images	
Long term preservation	High-resolution images of the sources are archived by the owning institutions.
Web presentation	For presentation on the web low resolution image formats are used (*.jpg, *.png); they are stored as data streams of the respective objects in the repository and presented using IIIF technology.
Audio	
Long term preservation	High-resolution audio files are archived by the owning institutions.
Web presentation	For presentation on the web compressed audio formats are used (*.mp3); they are stored as data streams of the respective objects in the repository.
Video	
Long term preservation	High-resolution video files are archived by the owning institutions.
Web presentation	For presentation on the web compressed video formats are used (*.mp4); they are stored as data streams of the respective objects in the repository.

Description of Metadata

Text	
Office documents Hierarchical textdata Program code	*Describe metadata and meta data formats here ...*
Images Long term preservation Web presentation	*Describe meta data and metadata formats here ...*
Audio Web presentation	*Describe meta data and metadata formats here ...*
Video Web presentation	

Data licensing

Project data is licensed either through Creative Commons licenses CC BY 3.0 AT[1] or CC BY-NC-SA 3.0 AT[2].

Documentation

The data generated in the project are continuously documented depending on the data type and are made available via the project website in the appropriate formats:

- TEI-P5 models through ODD-files. From these a verbal description is generated.
- *Describe documentation of other data models.*
- Data models that describe analysis strategies and display dependencies are represented graphically and described verbally. Data generated on this basis (TEI XML, RDF) will be described separately.
- Meta and annotation data are described verbally.
- The repository is described in detail at http://gams.uni-graz.at/o:gams.doku.

Ethical and legal aspects

- Conducting this project raises no ethical questions.
- In the course of the project copyrighted material will be used. The use of these

---

[1] https://creativecommons.org/licenses/by/3.0/at/deed.en
[2] https://creativecommons.org/licenses/by-nc-sa/3.0/at/deed.en

materials is either cleared with the owning parties who provided written letters of agreement or the material is used for demonstrative and explanation purposes according to §42f of the Austrian Copyright Law (Urheberrechtsgesetz).

Data Sharing

GAMS provides all standard export interfaces of a FEDORA-based OAIS-compliant asset management system. All data generated in the project are, corresponding to the licenses issued, available to research for free use.

Work Data

Data generated during the work process or before it is fed into the repository is stored at the dislocated workstations of the project staff and/or with select cloud services (Google, Dropbox, …) for better availability: this not only creates better collaboration environment, it also creates an external backup copy of the data. The data processed by staff at the ZIM-ACDH are stored locally on the respective workstations and via SVN-connection additionally on a network drive of the university NAS.

Repository

GAMS is an OAIS-compliant Asset Management System based on the Open Source software FEDORA and further developed by the ZIM-ACDH. The repository builds upon a web service-based (SOAP, REST), platform-independent and distributed system architecture, a largely XML based content strategy, the support of XML based import and export standards (METS, etc.) and the use of standardized data and metadata formats.

The Cirilo client, a java application developed for content preservation and data curation in FEDORA-based repository systems, includes object creation and management, versioning, normalization and standards, and the choice of data formats. In addition, standardized access methods via visualization tools have been realized. For example, METS/MODS objects were optimized for the DFG-Viewer format and TEI objects can be used as a basis for the implementation of the Voyant Tools or the Versioning Machine. User-defined mappings allow for a simple transformation of the most important content into RDF triples and consequently into Linked Open Data.

To ensure the long-term availability and reliable citability of the generated resources ZIM-ACDH is a member of the handle network runs its own handle server. All data objects in the system receive a persistent identifier (PID) based on handle.net and can thus be explicitly cited, analogous to a print publication.

Since 2014 GAMS has been a certified trusted digital repository in accordance with the guidelines of the Data Seal of Approval with a particular focus on the persistent storage and reusability of resources considered to be worthy of long-term preservation. GAMS is also registered with the Registry of Research Data Repositories re3data.org .

Reusability and interoperability of the produced research data is guaranteed by the use of adequate data standards (XML/TEI P5, METS/MODS in viewers, Dublin Core for the harvesting via standard OAI-PMH interfaces, jpg and TIFF for the facsimiles), standardized data models (RDF) and processing languages (XSLT).

Data Storage and Backup

Data storage for GAMS is provided via SAN by the University's IT department (UNI IT). Data is stored redundantly in two data centres in different campus buildings. Data backup in GAMS is part of the central backup processes of the University of Graz. Daily backups and are stored on a disk array and later moved to tape. There is an additional daily offsite backup managed by the Centre itself. The combination of both backups ensures accessibility of

backup data over a period of seven years.

Note:

This template is based on the work of the e-infrastructures austria project (http://e-infrastructures.univie.ac.at/). Further information on DMPs can be obtained at „Cluster C" under the following url: http://e-infrastructures.univie.ac.at/das-projekt/deliverables.

This DMP template is licensed through the CC BY license: https://creativecommons.org/licenses/by/3.0/at/deed.de