



## **Implementation of the Data Seal of Approval**

The Data Seal of Approval board hereby confirms that the Trusted Digital repository GESIS Data Archive for the Social Sciences 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 May 8, 2014.

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:	<a href="#">DSA-booklet_2014-2017.pdf</a>
All Guidelines Documentation:	<a href="#">Documentation</a>
Repository:	GESIS Data Archive for the Social Sciences
Seal Acquiry Date:	May. 08, 2014
For the latest version of the awarded DSA for this repository please visit our website:	<a href="http://assessment.datasealofapproval.org/seals/">http://assessment.datasealofapproval.org/seals/</a>
Previously Acquired Seals:	None
This repository is owned by:	<b>GESIS Data Archive for the Social Sciences</b> <ul style="list-style-type: none"><li>• Unter Sachsenhausen 6-8 50667 Cologne Germany</li></ul> <p>T +49 (0)221-47694-423 E Natascha.Schumann@gesis.org W <a href="http://www.gesis.org/">http://www.gesis.org/</a></p>

# Assessment

## 0. Repository Context

### Applicant Entry

*Self-assessment statement:*

GESIS – Leibniz-Institute for the Social Sciences is an infrastructure institution for the social sciences in Germany providing services at all stages of the research data lifecycle. The primary focus of the Data Archive, which is part of GESIS, is to provide excellent data service for national and international comparative surveys from the fields of social and political science research. The purpose of the Data Archive is to keep social science research data available for data sharing and re-use.

The Designated Community is constituted of researchers from the social sciences with a focus on empirical social research. Our collection policy defines the criteria a study should meet to be admitted into the Data Archive. The studies in the archive are described with standardized metadata by qualified staff. The metadata scheme used is DDI, a standard well established within the social sciences community. GESIS is also involved in the development and maintenance of the DDI standard. A detailed description is a precondition for data recovery as well as for ensuring correct re-use and interpretation for the future.

<http://www.gesis.org/en/home/>

All URLs accessed 07 April 2014

### 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:*

Sufficient information on depositing and archiving of social science research data is provided on the GESIS web pages. Remaining questions and details are clarified during the pre-ingest phase.

Pre-ingest communication with data producers adheres to important steps of the Preliminary and Formal Definition Phases as outlined in the Producer-Archive Interface Methodology Abstract Standard (PAIMAS) standard (CCSDS 2004). The Preliminary Phase is designed to gather necessary information, establish contact with data producers, and to provide the data producers with all necessary information – in particular about the archiving services offered by GESIS, the minimal requirements that have to be fulfilled for GESIS to accept data for archiving (see below), as well as licensing options and the archive agreement. During the Formal Definition Phase, agreements and schedules sketched before are discussed in more detail and on a more concrete level. In this phase, subject-related, technical, and legal questions are clarified, and agreements made with regard to delivery formats, schedules, etc. This phase concludes with the signing of the archive agreement. The Data Archive requests data producers to submit all materials necessary for a secondary analysis. This includes at least

- information about the primary researcher(s) and title of the study
- the data itself, prepared for direct use in statistical software packages if possible,
- the instrument or instruments used for data collection (e.g. questionnaire),

- a methodological description of the data collection and preparation procedures,
- publications or references to publications based on the respective data.

No materials are accepted that are subject to any copyright restrictions which may interfere with the use of the data as outlined in the archive agreement (e.g. copies of complete books).

Comprehensive (general) information on data management, data archiving, data sharing and re-use, as well as specific information on how to deposit with the GESIS Data Archive is given on the GESIS web pages. Beyond that GESIS actively promotes and supports all aspects of research data management, including e.g. long-term preservation, re-use, citation, through own research & development, publications, presentations, as well as active participation in a wide range of initiatives and projects.

GESIS makes broad information about a given study available to data consumers via web pages, data catalogues, variable reports and other publications so that secondary users can analyze the data and can assess its quality.

Links:

Data Archive Website: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/>

PAIMAS standard (CCSDS 2004): <http://public.ccsds.org/publications/archive/651x0m1.pdf>

Archive agreement:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/Archivierungsvertrag\\_GESIS\\_Datenarchiv\\_v9](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Archivierungsvertrag_GESIS_Datenarchiv_v9)

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

General Advice on Data Management & Archiving:

<http://www.gesis.org/en/archive-and-data-management-training-and-information-center/training-center-home/>

Jensen, Uwe (2012): Leitlinien zum Management von Forschungsdaten, Sozialwissenschaftliche Umfragedaten.

GESIS Technical Reports, 2012/07:

[http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2012/TechnicalReport\\_2012-07.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-07.pdf)

datorium: <https://datorium.gesis.org/xmlui/?locale-attribute=en>

Usage regulations: <http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/>

Data Catalogue and other portals: <http://www.gesis.org/en/services/research/>

Example of a study description (European Values Study 2008: Germany (EVS 2008)) in the Data Catalogue:

<https://dbk.gesis.org/dbksearch/SDesc2.asp?no=4753&tab=0&ll=10&notabs=&af=&nf=1&search=evs&search2=&db=E>

All URLs accessed 07 April 2014

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

## **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:*

Data are usually delivered in formats that are well-established within the social science community, in particular statistical formats like SPSS or Stata. The Data Archive has published a list of recommended formats on its website.

After a submission has been received, the data and all accompanying material are assessed with regard to content, structure and format. Incoming objects in non-compliant formats will be converted to defined archival formats suitable for long-term preservation. If that is not possible objects resp. studies are not accepted for archiving or, exceptionally, bit-level preservation is offered. However, to provide users with more comprehensive information on this matter we have published a list of preferred formats.

For each submission, archive staff checks whether the delivered material is complete, correct, and in a suitable technical condition (e.g. readable, virus free, etc.). Further checks concerning the plausibility, consistency, data weighting and data protection are carried out. This ingest control is carried out based on an internal checklist which at same time is used for documentation purposes.

For producing long-time preservation data, we currently employ StatTransfer. However, to obtain more control over this process presently a tool is being developed. At the time being, it is limited to handling .sav files, but shall be expanded to cover more data formats (Stata coming next). It exports the table data to a tab separated file and the metadata (variable names, labels, format specifications, value labels...) to various formats including an in-house intermediary format, DDI-2.5, SPSS syntax files, and sql database definitions.

To ensure data consistency, redundant exports from various tools are being mapped onto each other and transformed according to specific tools' bugs and limitations.

### **Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

Extensive testing is done to ensure proper codepage handling (incoming from utf-8, windows-1252, iso-8859-2, shift\_jis, DIN\_66003) and all archival data is written in utf-8.

All non-trivial changes in the dataset (rounding from -> to, codepage conversion from -> to...) are logged and stored away, to allow full debugging of the conversion process.

Links:

Data Archive Website: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/>

Recommended formats:

<http://www.gesis.org/en/services/archiving-and-registering/data-archiving/preparing-data-for-submission/>

All URLs accessed 07 April 2014

## **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:*

Most of the metadata are created by archive staff on the basis of the (usually un-structured) documentation delivered by the depositors. Under certain circumstances (e.g. the deposit consists of several studies/ data sets) data depositors are asked to fill in a simple off-line study description template. It requests information such as "Title", "Alternative Title", "Date of data collection" and some others. If the metadata are not sufficient, an attempt is made to extract it from publications. Otherwise data depositors are asked to deliver additional information.

A web-based deposit form is currently only available in connection with a recently introduced data sharing platform (datorium) which serves as a low-threshold opportunity for individual researchers and projects to archive and share their data. The deposit form developed for this service will also be used in the future for regular deposits to the archive.

The metadata scheme used by the Data Archive is compliant with the DDI Standard, as well as with the da|ra and DataCite metadata schemes. Resource discovery metadata is available for download in DDI2 and DDI3 format under a CC0 license. Datasets of particular importance and important study collections falling into the Data Archive's core areas of collection are processed (cumulated, harmonized, standardized), documented, and enhanced in much greater depth – not only on study level, but on the level of individual questions and variables. This metadata is used for different purposes (e.g. production of codebooks/ variable reports; long-term preservation) and is made available online via portals like ZACAT. The deployed tools (e.g. Dataset Documentation Manager, CodebookExplorer) produce DDI compliant metadata as well.

Further structural and administrative metadata is created for internal use. Among others, this provides relevant technical and provenance information. Detailed information on the metadata scheme was published in Zenk-Möltgen and Habel 2012 (in German).

The indexing and documentation of data in studies/study collections is carried out using national and international standards, and classifications where necessary and relevant. These include:

- ISO 3166 - Codes for the representation of names of countries and their subdivision
- ISO 639 - Codes for the representation of names of languages
- ISCO - International Standard Classification of Occupations
- NUTS - Nomenclature of territorial units for statistics
- Description of study content:
- ZA Classification  
(<http://www.gesis.org/unser-angebot/recherchieren/thesauri-und-klassifikationen/zentralarchiv-klassifikationsschema-za-klassifikation>)
- CESSDA Topic Classification  
(<http://www.gesis.org/unser-angebot/recherchieren/thesauri-und-klassifikationen/cessda-topic-classification/#c28449>).
- In addition, the European Language Social Science Thesaurus (ELSST;  
<http://www.gesis.org/unser-angebot/recherchieren/thesauri-und-klassifikationen/european-language-social-science-thesaurus-elsst/#c2>) makes the content of the ZACAT Online Catalog browsable by keyword in the CESSDA portal.

Links:

datorium: <https://datorium.gesis.org/xmlui/?locale-attribute=en>

Linne, Monika (2013): Sustainable data preservation using datorium: facilitating the scientific ideal of data sharing in the social sciences. In: Borbinha, José; Nelson, Michael; Knight, Steve (Hrsg.): Proceedings of the 10th International Conference on Preservation of Digital Objects, Lisbon: Biblioteca Nacional de Portugal, S. 150-155: [http://purl.pt/24107/1/iPres2013\\_PDF/Sustainable%20Data%20Preservation%20using%20datorium%20%E2%80%93%20facilitating%20](http://purl.pt/24107/1/iPres2013_PDF/Sustainable%20Data%20Preservation%20using%20datorium%20%E2%80%93%20facilitating%20)

DDI Codebook: <http://www.ddialliance.org/Specification/DDI-Codebook/>

da|ra Metadata Schema: <http://dx.doi.org/10.4232/10.mdsdoc.2.2.1>

DataCite Metadata Schema: <http://schema.datacite.org/>

DBKfree: <http://www.gesis.org/en/services/research/data-catalogue/dbkfree/>

ZACAT: <http://zocat.gesis.org/webview/>

Dataset Documentation Manager and CodebookExplorer:  
<http://www.ddialliance.org/node/856>  
<http://www.gesis.org/en/services/research/codebookexplorer/>  
<http://info1.gesis.org/projects/Default.aspx?alias=info1.gesis.org/projects/dsdm>

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

Zenk-Möltgen, Wolfgang; Habel, Norma (2012): Der GESIS Datenbestandskatalog und sein Metadatenchema, Version 1.8. GESIS Technical Reports, 2012/01. (in German):

[http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2012/TechnicalReport\\_2012-01.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-01.pdf)

Zenk-Möltgen, Wolfgang: Metadata in the Data Catalogue DBK at the GESIS Data Archive:

[http://www.ratswd.de/ver/docs\\_PID\\_2012/Zenk-Moeltgen\\_PID2012.pdf](http://www.ratswd.de/ver/docs_PID_2012/Zenk-Moeltgen_PID2012.pdf)

All URLs accessed 07 April 2014

## **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:*

According to GESIS's by-laws, among the association's primary objectives is the "archiving, documentation, and long-term preservation of social sciences data, including the indexing of data as well as the high-quality enhancement of particularly relevant data to prepare them for re-use" (§ 2). Thus, GESIS voices its commitment to preserve and provide access to social sciences research data in its by-laws and mission statement.

The Data Archive's preservation principles and practices have also been communicated in contributions to relevant publications (e.g. Quandt, Mauer 2012; Jensen, Katsanidou, Zenk-Möltgen 2011; Jensen 2012).

In the past, activities of the archive mainly concentrated on domain-specific aspects: understanding and curating social science research data. As a consequence the archive has strong expertise in curating empirical social science research data and is well integrated into a network of European and international social science archives. In addition to this traditional expertise in curating data, the Data Archive has recently focused more strongly on long-term preservation aspects. Today, GESIS is involved in initiatives dealing with topics such as persistent identifiers, metadata standards, building of data infrastructures etc. Since 2013 GESIS is partner of nestor, the German competence network for digital preservation. All these activities contain promotional aspects as well.

We recently formulated a preservation policy, which describes the main principles of digital preservation activities within our archive. The activities of the archive are documented on our website. GESIS is actively communicating its services and resources in numerous ways (exhibition stands at conferences, social media, publications, web pages, brochures etc.). The Archive and Data Management Training Center explicitly promotes research data management and data curation in the social sciences. It offers workshop and training events in data management for researchers and in digital preservation for archive and repository staff.

We have several (mostly internal) strategic and planning documents which outline how the mission statement is implemented. The annual report covers research and research based services from all departments of GESIS. Program budget and long-term planning are further instruments for strategic issues.

Currently no formal succession plan is in place. This issue needs to be addressed in the future. However, as the largest infrastructure for the social sciences in Germany, GESIS operates under a relatively stable financial

framework. Due to the importance of the institution for Germany, it is jointly funded by the Federation and the Länder.

Links:

GESIS Constitution (in German): <http://www.gesis.org/das-institut/der-verein/satzung/>

GESIS Mission Statement: <http://www.gesis.org/en/institute/the-association/mission/>

Quandt, Markus; Mauer, Reiner (2012): Sozialwissenschaften. In: Neuroth, Heike; et al: Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme, Boizenburg: vwh, S. 61-8:  
[http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/kapitel/nestor\\_bestandsaufnahme\\_004.pdf](http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/kapitel/nestor_bestandsaufnahme_004.pdf)

Neuroth, Heike; et al: Digital Curation of Research Data: Experiences of a Baseline Study in Germany  
<http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/index.php?lang=en>

Jensen, Uwe; Katsanidou, Alexia; Zenk-Möltgen, Wolfgang (2011): Metadaten und Standards. In: Büttner, Stephan; et al.: Handbuch Forschungsdatenmanagement, Bad Honnef: Bock u. Herchen, S. 83-100:  
<http://opus4.kobv.de/opus4-fhpotsdam/frontdoor/index/index/docId/198>

Jensen, Uwe (2012): Leitlinien zum Management von Forschungsdaten, Sozialwissenschaftliche Umfragedaten. GESIS Technical Reports, 2012/07,  
[http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2012/TechnicalReport\\_2012-07.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-07.pdf)

nestor: [http://www.langzeitarchivierung.de/Subsites/nestor/EN/Home/home\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/EN/Home/home_node.html)

Preservation Policy:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/DAS\\_Preservation\\_Policy\\_eng.pdf](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/DAS_Preservation_Policy_eng.pdf)

Archive and Data Management Training Center:

<http://www.gesis.org/en/archive-and-data-management-training-and-information-center/training-center-home/>

Annual reports (in German): <http://www.gesis.org/das-institut/presse-und-informationen/informationsmaterial/>

All URLs accessed 07 April 2014

## **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:*

GESIS – Leibniz Institute for the Social Sciences is legally registered as a non-profit association and is sponsored jointly by the federal government and the federal states according to Article 91b of the German Federal Constitution. Funding in accordance with Article 91b is assigned in seven-year cycles. As a member of the Leibniz Association, the umbrella organization to currently 87 research institutions which “conduct research and provide infrastructure for science and research and perform research-based services – liaison, consultation, transfer – for the public, policy-makers, academia and business” (<http://www.research-in-germany.de/leibniz>), GESIS is part of a strong network of publicly funded research institutions. In addition, GESIS has long-lasting and strong ties with universities. GESIS has three partner universities (GESIS’ president and heads of departments are professors at these universities) and 60 German universities are members of GESIS e.V.

To enable the Data Archive to preserve and offer data for re-use, data producers sign an archive agreement when submitting data for archiving. Through this archive agreement, the archive receives all necessary rights of (non-exclusive) use as laid down in German copyright law (especially §§16 and 19 UrhG). Thus the Data Archive receives permission from the data producers to carry out long-term preservation actions, e.g. migration to a different file format, as well as making several copies of the data and their documentation for backup and distribution. In addition the data producer assures that legal obligations regarding personal or confidential data are adhered to. Furthermore data producers determine under which standard licensing conditions the data is made available to end users. They can choose between different levels, ranging from open access for every registered user up to restricted access where access is only granted after written permission from the data producer (in some cases special contracts additionally need to be signed by data users).

All incoming data is checked for disclosure risks when submitted to the archive. If necessary, data is anonymised in cooperation with the data producer. Data with disclosure risk is only made available under special conditions.

Within the Data Archive two specially trained staff members are responsible for the management of data with disclosure risks. Furthermore GESIS has an external data protection officer who is regularly involved in all relevant issues. Additionally, GESIS consults lawyers for advice and expert opinion in cases where further legal

### **Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)



advice is required.

The user contract for the use of confidential data contains different sanctions to apply in case of non-compliance: In case of misuse the user has to delete all data and supplementary material. In addition, a report will be sent to other data service centers as well as to the German Data Forum (RatSWD). It is also possible that the data user's account is blocked temporarily or permanently.

The Secure Data Center (SDC) provides controlled and secure access to data deserving special protection. Data protection legislation requires that the possibility of re-identifying individuals in the data provided by GESIS must be avoided. The SDC offers restricted access to data which has not been fully anonymised. It uses special contracts and user guidelines which are published on the SDC website. The SDC applies special procedures to manage data with disclosure risk, such as carefully vetting researchers, requiring the signing of an agreement and applying various measures of organisational and technical control to protect data. The service is not yet fully set up and under development, so while the first users had already been granted access to the data, further steps are now being taken to improve procedures for secure data handling, better training of staff with regards to secure data handling and the consequences if users do not comply with the agreed procedures.

Links:

Leibniz Association: <http://www.research-in-germany.de/leibniz>

Archive agreement:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/Archivierungsvertrag\\_GESIS\\_Datenarchiv\\_v9](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Archivierungsvertrag_GESIS_Datenarchiv_v9)

Terms of use/ usage regulations:

<http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/>

Charge regulations: <http://www.gesis.org/en/services/data-analysis/data-archive-service/charges/>

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

Contract for on-site use:

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_SDC\\_Data\\_Use\\_Agreement\\_On\\_Site.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_SDC_Data_Use_Agreement_On_Site.pdf)

Contract for off-site use:

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_SDC\\_Data\\_Use\\_Agreement\\_Off\\_Site.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_SDC_Data_Use_Agreement_Off_Site.pdf)

German Data Forum: <http://ratswd.de/en>

GESIS – Secure Data Center:

<http://www.gesis.org/en/services/data-analysis/data-archive-service/secure-data-center-sdc/>

Technical and organizational measures (in German):

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_Hinweise\\_zum\\_SDC\\_Gastarbeitsplatz.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_Hinweise_zum_SDC_Gastarbeitsplatz.pdf)

Security requirements for off-site access (in German):

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_Merkblatt\\_Umgang\\_Forschungsdaten.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_Merkblatt_Umgang_Forschungsdaten.pdf)

All URLs accessed 07 April 2014

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

## **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:*

The main principles of long-term preservation as carried out by the GESIS Data Archive are documented in our preservation policy.

The security and risk management is carried out in close co-operation with GESIS's IT department, which administers the servers and takes care of backups, media monitoring and refreshing.

To protect the data, the following backup and access control procedures are in place to guarantee the (physical) safety of the digital archive holdings:

1) Physical protection measures:

a. The computing center and server rooms are secured against unauthorized access by means of an electronic access control system.

b. Smoke and water detectors are in place, temperatures in the computing center are monitored.

2) Redundant data storage in different locations (Cologne and Mannheim):

a. Frequent (up to daily) incremental and complete back-ups to onsite disk and tape libraries (tapes stored in suitable vault). In addition, frequent backups to offsite tape libraries.

3) Diversity of storage media (hard disk, tape) and frequent media refreshment.

The backup and storage procedures (redundant and distributed storage) in place allow for fast and complete recovery / restoral of the archive holdings in case of a disaster.

In addition to the backup procedures described above, the Archive has the following technological and organizational measures in place to assure that the data bitstream is securely stored and cannot be altered without authorization: Write access to the archive server is highly restricted and governed by a set of strict rules and regulations. Only two members of the Data Archive staff are authorized to add, delete or change files on the

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

archive server. The transfer of files into the archive takes place by means of a special transfer folder in the network, from which data and documentation to be archived are picked up, checked once again for conformity with the Data Archive's preservation standards (file formats, naming conventions, etc.), and transferred onto the archive server by the two authorized staff members.

All source and archive files' checksums are stored to ensure data consistency. All checksumming is done using two checksums (currently sha1 and sha512), to allow seamless migration to a new checksum mechanism and to handle the unlikely event of checksum collisions.

Links:

Preservation Policy:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/DAS\\_Preservation\\_Policy\\_eng.pdf](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/DAS_Preservation_Policy_eng.pdf)

Data Service Infrastructure for the Social Sciences and Humanities. "Report about Preservation Service Offers", Deliverable: D4.2, p 47-70:

[http://dasish.eu/publications/projectreports/D4.2\\_-\\_Report\\_about\\_Preservation\\_Service\\_Offers.pdf](http://dasish.eu/publications/projectreports/D4.2_-_Report_about_Preservation_Service_Offers.pdf)

All URLs accessed 07 April 2014

## **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:*

An important task and the basis for all preservation activities is the validation of the submitted data. In addition GESIS has implemented a versioning policy (including documentation of all changes made to the data) and assigns persistent identifiers to each version of a study.

The first activity within the ingest process is to check whether all delivered material is complete, correct and in a suitable technical condition (e.g. readable, virus free, etc.). Further checks concerning plausibility, consistency, data weighting and data protection are carried out.

Undertaking extensive validation action during the ingest phase enables us to gain an exact overview about the data we have in the Archive. We know the formats and can estimate the risks in terms of preservation. To avoid future obsolescence, we accept only those formats we know we are capable of preserving. Data submitted in other formats are converted during the ingest into suitable ones. If data is in a format we cannot handle or convert, we direct data producers to a more suitable archive.

Another important part of our approach is the creation of standardized metadata. Besides technical and administrative information, providing extensive context information is essential for future usability and interpretability of social science research data.

Supported by a constant monitoring of technology (storage technology and media, software and file formats) as well as a normalization of file formats during ingest, the Data Archive pursues a migration strategy to ensure long-term access to its holdings. Data and documentation are archived in well-defined, standardized file formats to ensure that efficient migration strategies can be developed when this becomes necessary. Syntax/ setup files documenting the changes between different versions are kept in addition.

While the refreshing of storage media takes place continuously, format migrations are undertaken only if the readability / interpretability of archive holdings is endangered by technological obsolescence, if they cannot be

processed and used anymore in a state-of-the-art manner, and/or if a format migration brings considerable advantages with regard to user-friendliness and the work of the archive.

During format migrations utmost care is taken not to alter the significant properties of the archival objects. Being domain experts, all central aspects of social science data (e.g. characteristics of data matrices and variable definitions) and the special features of the different file formats are well-known to archive staff. However, we are currently approaching the topic of significant properties in a more systematic way with the objective of developing a detailed and testable definition (see presentation at Open Repositories 2014).

Migration procedures are documented thoroughly in order not to compromise the archival objects' authenticity during the migration process.

Existing internal documentation is available to all staff members through a wiki.

Links:

Data Archive Webpage, Ingest: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/ingest/>

Preservation Policy:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/DAS\\_Preservation\\_Policy\\_eng.pdf](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/DAS_Preservation_Policy_eng.pdf)

Data Service Infrastructure for the Social Sciences and Humanities. "Report about Preservation Service Offers", Deliverable: D4.2, p 47-70:

[http://dasish.eu/publications/projectreports/D4.2\\_-\\_Report\\_about\\_Preservation\\_Service\\_Offers.pdf](http://dasish.eu/publications/projectreports/D4.2_-_Report_about_Preservation_Service_Offers.pdf)

Mauer, Reiner (2012): Das GESIS Datenarchiv für Sozialwissenschaften. In: Altenhöner, Reinhard; Oellers, Claudia (Hrsg.): Langzeitarchivierung von Forschungsdaten. Standards und disziplinspezifische Lösungen, Berlin: Scivero, S. 197-215, [http://ratswd.de/dl/downloads/langzeitarchivierung\\_von\\_forschungsdaten.pdf](http://ratswd.de/dl/downloads/langzeitarchivierung_von_forschungsdaten.pdf)

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

Astrid Recker, Stefan Müller: A rose is a rose is a rose? Defining significant properties of social science research data GESIS - Leibniz Institute for the Social Sciences, Germany (Presentation at the Open Repositories):  
[https://www.conftool.com/or2014/index.php?page=browseSessions&form\\_session=31](https://www.conftool.com/or2014/index.php?page=browseSessions&form_session=31)

All URLs accessed 07 April 2014

## **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:*

Empirical research projects producing research data generally go through a process of multiple phases – the research data cycle. Each individual phase of this cycle requires specific know-how to obtain significant results. GESIS has the know-how and optimizes utilization of this by offering a unique package of services to accompany the entire research data cycle. Archiving is well integrated into this data life cycle. Workflows within the archive are organised according to an archival life cycle, ranging from acquisition/ pre-ingest to dissemination of data. The central functions of the OAIS reference model can be mapped to the existing structure of the archive. For most parts of the corresponding workflows, procedures, standards and rules are in place. Even though internal documentation exists it currently is not complete and up to date for every activity. Significant steps to improve this situation were already taken, but there is still some work to be done.

An overview about the different steps and processes (pre-ingest, ingest, study description, archiving, access) can be found on our website.

All transformations made to data are documented. All significant corrections/changes of the data will be discussed with data depositors beforehand. General information about handling of the data is given on our website and as well during pre-ingest communication with depositors. Standard procedures for changes applied to the data during ingest or at later stages are available (e.g. naming conventions, handling of missing data, versioning rules). Existing internal documentation is available to all staff members through a wiki. Most of the staff members in the archive hold degrees in social sciences and thus have an excellent understanding of the data they are curating. In addition, some members of the Data Archive hold degrees in library and information sciences. Staff members regularly take part in internal and external trainings on data management, metadata, long-term preservation and other relevant fields.

The GESIS Data Archive is specialized on quantitative survey data, but holds other types (e.g. aggregate/ time series data or textual materials) as well. Since other types of data (e.g. transactional data, social media data, and experimental data) are becoming more and more important the archive evaluates opportunities and challenges with regard to archiving and re-use and is also engaged in corresponding projects.

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)



According to the GESIS by-laws, GESIS's service and data offers are to enable and support research of societal developments in national and international comparative and historical perspective. According to the Data Archive's acquisition policy, data are therefore not collected with a tight-knit thematic focus, but rather with the objective of building a broad collection of data of general relevance to social sciences research. Thus, while there are certain thematic areas in which the Data Archive has built particularly comprehensive holdings, the collection policy does not per se limit the thematic focus of data to certain areas.

The Data Archive acquires data both actively (i.e. by contacting relevant projects, surveys, or programs) and passively (i.e. by addressing potential depositors through the website or publicity materials). Data are evaluated by means of criteria enabling us to determine their relevance and their suitability for archiving by GESIS. In cases where archiving with GESIS cannot be realized, we try to find alternative data centres (e.g. partners in the network of research data centres of the German Data Forum) and bring data producers into contact with them. If there is no alternative option, GESIS tries to act as (temporary) fallback option.

The privacy of research subjects is guaranteed through different measures: When ingested, all data is checked for disclosure risks and if necessary appropriate actions are taken (e.g. anonymisation; access is granted only in appropriate manner, i.e. through our Secure Data Center or with a special user contract). Trained staff is available who can give advice on handling and processing of sensitive data.

Links:

Data Archive Website: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/>

Schumann, Natascha; Mauer, Reiner: The GESIS Data Archive for the Social Sciences: A Widely Recognised Data Archive on its Way. The International Journal of Digital Curation. Vol 8, No 2 (2013).  
<http://dx.doi.org/10.2218/ijdc.v8i2.285>

Collection Policy (in German):

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/2013-07-08\\_Was\\_wir\\_sammeln.pdf](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/2013-07-08_Was_wir_sammeln.pdf)

Poster CharmStats: [http://prezi.com/yj0x3ditpui\\_/charmstats-quick-harmonization/](http://prezi.com/yj0x3ditpui_/charmstats-quick-harmonization/)

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

All URLs accessed 07 April 2014

## **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:*

GESIS's offers are primarily directed at researchers (both in universities and non-university research institutions) and students – in particular in empirical social research with a focus on the areas of sociology and political science as well as social science in its entirety. The Data Archive strongly promotes data sharing and re-use and hence seeks to make data available as openly and easily accessible as possible. However, both legal regulations and respect for the needs and requirements of data depositors make it necessary to manage access according to principles laid out in our usage regulations.

To enable the Data Archive to preserve and offer data for re-use, data producers must sign an archive agreement when submitting data for archiving. According to the archive agreement, the Data Archive may archive all data and documentation and process them further for the purpose of long-term preservation and re-use. The archive receives all necessary rights of (non-exclusive) use as laid down in German copyright law (especially §§16 and 19 UrhG). Thus the Data Archive receives permission from the data producers to carry out long-term preservation actions, e.g. migration to a different file format, as well as making several copies of the data and their documentation for backup and distribution.

In the archive agreement the data producers also determine under which standard licensing conditions the data will be made available to data users. Typically data producers choose to make data and documents available for scientific analysis carried out in academic research and teaching.

Currently there is no dedicated crisis management plan in place. However, data and all relevant access systems are redundantly stored and sufficiently documented to rebuild the repository.

Links:

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

Archive agreement:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/Archivierungsvertrag\\_GESIS\\_Datenarchiv\\_v9](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Archivierungsvertrag_GESIS_Datenarchiv_v9)

Terms of use/ usage regulations (incl. access conditions):

<http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/>

All URLs accessed 07 April 2014

### **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 data is offered in formats that are well-established in the community, such as SPSS, Stata and other statistical formats. The DIPs are normally available in one or more of these formats. Other formats can be produced on request. All studies are described in and are accessible through our Data Catalogue. As stated in GESIS's by-laws, among its primary tasks is "to create user-friendly and high-quality possibilities for access to all the information and data relevant to empirical social research [...]" (§2e). To facilitate access to the data, all data sets that can be shared without further consultation with the data depositor are available for download via different platforms. Users are required to register once and log-in before downloading a data set. As the registration and download system are fully automated, the data are available free-of-charge immediately and at any time convenient to users.

The data catalogue is the central access point to the holdings of the archive and comprises of study descriptions for all archived studies, mainly including micro data from survey research and aggregate time series data. Further portals allow for access to special holdings and/or offer particular additional services, e.g. the ZACAT Online Study catalogue, which provides selected studies with extensive documentation on study and variable level, or HISTAT, a system giving access to historical time series data.

In addition to the central systems for immediate download, users can order data in access categories B and C from the Archive's Data Service via a shopping cart system, by e-mail or telephone. They will receive this data (in customized form, if they wish) on a CD-ROM or DVD or via a secure download using Cryptshare. For this data service handling fees are charged.

Different tools to search, explore, and analyze studies (data sets and accompanying material) are in place (partly allowing searches down to the level of individual variables/ questions). Harvesting of metadata of our holdings published via ZACAT is in place. An OAI interface (OAI-PMH) for our data catalogue (DBK) is planned to be implemented during this year. All metadata from GESIS DBK are available under the Creative Commons CC0 1.0 Universal Public Domain Dedication. However, GESIS requests that users actively acknowledge and give attribution to all metadata sources, such as the data providers and any data aggregators, including GESIS.

Each study is assigned a Digital Object Identifier (DOI), a permanent, persistent identifier used for citing and linking electronic resources. GESIS not only assigns DOIs to its own holdings but as a member of the Data Cite Consortium and maintainer of the allocation agency data also offers DOI services to a wider national and international community.

Links:

Data Catalogue: <http://www.gesis.org/en/services/research/data-catalogue/>

Charges: <http://www.gesis.org/en/services/data-analysis/data-archive-service/charges/>

Access categories:

[http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/#3\\_Access\\_categories](http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/#3_Access_categories)

ZACAT Online Study Catalogue: <http://www.gesis.org/unser-angebot/recherchieren/zacat-online-study-catalogue/>

Histat (in German): <http://www.gesis.org/en/services/data-analysis/daten-historischer-studien/histat/>

sowiport: <http://www.gesis.org/sowiport>

SIMon - Social Indicators Monitor: <http://www.gesis.org/simon>

MISSY: <http://www.gesis.org/missy/en/missy-home/>

Usage Guidelines for GESIS DBK Metadata: <https://dbk.gesis.org/dbksearch/guidelines.asp?db=e>

DOI: <http://www.doi.org/>

da|ra: <http://www.da-ra.de/en/home/>

All URLs accessed 07 April 2014

## **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:*

The integrity of data is monitored with the help of checksums among other measures. All checksumming is done using two checksums (currently sha1 and sha512), to allow seamless migration to a new checksum mechanisms and to handle the unlikely event of checksum collisions.

Automated scripts generate and compare checksums of all objects in the central archive storage. Output is logged and checked by staff regularly. The central metadata management system (DBKEdit) is accessible to authorized staff only and all activities and changes are logged.

Each study receives a version number. The versioning of data is governed by a versioning guideline that is strictly adhered to in order to meet the requirements of DOI-assignment and the standards of trusted digital preservation. The version number is a three-digit number (major.minor.revision). The major number is incremented when there are changes in the composition of the data set (e.g. additional variables or cases), the minor or second number is incremented when significant errors have been fixed (e.g. coding errors, misleading value labels), and the third or revision number is incremented when minor bugs are fixed (e.g. spelling errors in variable or value labels). The version number is included as a variable in the data set and added to file names of corresponding objects in accordance with a set of naming conventions. The version history, indicating the major/minor changes made to the data, is documented in the metadata management system and made available to end users through our data catalogue (DBK). All data set versions – from the original to the latest version – are kept in the archive. Syntax/setup files documenting the changes between the different versions are kept in addition.

Links:

Versioning guidelines: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/ingest/>

All URLs accessed 07 April 2014



## **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:*

After a submission has been received, the data and all accompanying material are assessed with regard to content, structure and format. After this initial completeness check on the basis of an internal checklist and technical control, all files are saved on the archive server in their original versions and formats – this SIP will not be altered anymore and will be retained in its original form. Information about depositor, archive agreement, deposition (data, responsible staff, composition of the SIP etc.) is compiled and documented. Corresponding checksums will be produced and stored. Subsequently, the data are converted to or saved to archival formats and undergo further checks and – where necessary – corrections. Any corrections carried out at this stage are documented in a way that allows for reverting to the original version at any time. All significant corrections/changes of the data will be discussed with data depositors beforehand. All changes are documented in syntax/ setup or further documentation files. These always contain additional information about by whom, when and why changes were made. Handling of different versions is supported by a versioning policy (see answer Guideline 11).

Identity of depositors is not checked in a formalized way but as a rule the staff has personal communication with every depositor (e-mail, telephone or face to face). Data depositors need to sign a archive agreement. In many cases longstanding relationships exist.

Links:

Versioning guidelines: <http://www.gesis.org/en/services/archiving-and-registering/data-archiving/ingest/>

Archive agreement:

[http://www.gesis.org/fileadmin/upload/institut/wiss\\_arbeitsbereiche/datenarchiv\\_analyse/Archivierungsvertrag\\_GESIS\\_Datenarchiv\\_v9](http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Archivierungsvertrag_GESIS_Datenarchiv_v9)

All URLs accessed 07 April 2014

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

## **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:*

Most of the workflows of the archive have grown over decades. But nonetheless, the archiving workflows conform to the OAIS reference model (see also answer Guideline 8). The main functional entities, Ingest, Archival storage, Data management, Administration, Preservation Planning and Access are implemented to considerable extent. We also have a Pre-ingest step, as described in the PAIMAS standard. The Pre-ingest includes acquisition activities as well as negotiations how data should be delivered and which material the Dissemination Information Package (DIP) should contain.

To gain a more structured overview of workflows and processes and to identify and close possible gaps a mapping between the Archive and the OAIS functional model, as well as an application of the concepts from the OAIS information model has been carried out for the most parts. Further conceptual work is on the way and concentrates on relevant aspects for our archival work, e.g. with regard to ingest processes (see conference programme Archiving 2014). Therefore, a complete mapping is not published yet. Even though not all processes are carried out exactly in the way defined by the OAIS reference model, the GESIS Data Archive fulfils all responsibilities described in OAIS reference model.

Production, collection, management and exchange of metadata is essential to research data archiving and management. The metadata schemes used by the Data Archive are compatible to DDI, which is an international specification for describing data from the social, behavioral, and economic sciences. The reference point for the technical infrastructure employed and developed by GESIS for data management and archiving is DDI. Expressed in XML, the DDI metadata specification supports the entire research data life cycle. DDI metadata accompanies and enables data conceptualization, collection, processing, distribution, discovery, analysis, repurposing, and archiving.

The published metadata for a dataset on study level is available for download in DDI2 and DDI3 format. Datasets of particular importance and important study collections falling into the Data Archive's core areas of collection are processed (cumulated, harmonized, standardized), documented, and enhanced in much greater depth – not only on study level, but on the level of individual questions and variables. The deployed tools also produce DDI compliant metadata.

Services offered by the Archive and the corresponding underlying infrastructure are constantly evaluated either as part of internal procedures (e.g. portfolio analysis, reports to and evaluations by the scientific advisory board and the user advisory board) or through external evaluation (Leibniz institutes are evaluated every 7 years). The current priorities for the development of the Archive are focused on professionalization in the area of preservation, handling of sensitive data, integrating new forms of data (e.g. social media data, georeferenced data), development of collaborative archiving solutions (as well in national as in international perspective), definition of services and as a general topic, the improvement of internal procedures and the integration of central internal production tools (project STARDAT). Regarding GESIS at large, the focus of infrastructure development is the buildup of an integrated technical architecture. Another task is the integration of different information services for end users offered in the areas of methods, current research information and literature, data management and archiving.

Links:

Recker, Astrid; Schumann, Natascha (2013): De-mystifying OAIS compliance. benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive. In: IASSIST Quarterly, Jg. 36, H. 2, pp. 6-11:

<http://www.iassistdata.org/iq/de-mystifying-oais-compliance-benefits-and-challenges-mapping-oais-reference-model-gegis-data-arc>

Recker, Astrid; Schumann, Natascha (2013): De-mystifying OAIS compliance: benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive. IASSIST 2013 "Data Innovation: Increasing Accessibility, Visibility and Sustainability". Köln, 28.05.-31.05. 2013.:

<http://de.slideshare.net/GESISarchivetraining/iassist-a5-reckerschumannslideshare>

Recker, Astrid; Schumann, Natascha (2014): Evaluating and Developing Ingest Workflows with the OAIS Reference Model

[http://www.imaging.org/ist/conferences/archiving/Archiving2014\\_Preliminary\\_Program.pdf](http://www.imaging.org/ist/conferences/archiving/Archiving2014_Preliminary_Program.pdf)

DBKfree: <http://www.gegis.org/en/services/research/data-catalogue/dbkfree/>

Dataset Documentation Manager: <http://www.ddalliance.org/node/856>

Zenk-Möltgen, Wolfgang; Habel, Norma (2012): Der GESIS Datenbestandskatalog und sein Metadatenschema,

Version 1.8. GESIS Technical Reports, 2012/01. (in German):

[http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2012/TechnicalReport\\_2012-01.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-01.pdf)

Zenk-Möltken, Wolfgang (2013): Utilizing DDI-lifecycle in the STARDAT project to manage data documentation. IASSIST 2013 - 39th Annual Conference "Data Innovation: Increasing Accessibility, Visibility and Sustainability". Cologne, 28.05.-31.05. 2013:

[http://www.iassistdata.org/downloads/2013/2013\\_c3\\_zenk-moltgen.pdf](http://www.iassistdata.org/downloads/2013/2013_c3_zenk-moltgen.pdf)

Mühlbauer, Alexander (2011): Creating a Prototype Application Compatible with DDI 3.1 for the STARDAT Project. 3rd Annual European DDI Users Group Meeting. Göteborg, 05.-06.12. 2011:

[http://www.iza.org/conference\\_files/EDDI2011/EDDI11%20D2%20-%20Alexander%20Muehlbauer%20-%20Creating%20a%20Prototy](http://www.iza.org/conference_files/EDDI2011/EDDI11%20D2%20-%20Alexander%20Muehlbauer%20-%20Creating%20a%20Prototy)

All URLs accessed 07 April 2014

## 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:*

The data deposited with the Data Archive are distributed through various channels. To facilitate access to the data, all data sets without special access restrictions are available for download via different platforms. Users are required to register once (by doing so accepting the terms of use/usage regulations) and log-in before downloading a data set. These data sets are available free-of-charge and at any time convenient to users.

Data belonging to access categories B and C (written permission through the depositor is needed for granting access), can be ordered via a shopping cart system, by e-mail or telephone. Users will receive data (in customized form, if they wish) on a CD-ROM or DVD or via a secure download using Cryptshare. For this data service small handling fees are charged.

Different access categories determined by the data producer and fixed in the archiving contract define the way data can be accessed and used. The conditions are outlined in the usage regulations as well as in the charge regulations.

The usage regulations also contain general access conditions. Users have to agree that they will inform the Data Archive when their project is completed, quote all used documents according to scientific conventions and to send two specimen copies of their publication to the Data Archive.

The use of metadata from GESIS DBK is possible under a Creative Commons licence. All metadata are available free of restriction under the Creative Commons CC0 1.0 Universal Public Domain Dedication. However, GESIS requests that users actively acknowledge and give attribution to all metadata sources, such as the data providers and any data aggregators, including GESIS.

Users of the recently introduced data sharing platform datorium, which serves as a low-threshold opportunity for individual researchers and projects to archive and share their data, are advised to use standardized licences like Creative Commons or Open Data Commons.

Access to confidential data is only given in our Secure Data Center. It provides controlled and secure access to data deserving special protection. Data protection legislation requires that the possibility of re-identifying individuals in data provided by GESIS must be avoided. The SDC offers restricted access to data which has not been fully anonymised. It uses special contracts and user guidelines which are published on the SDC website. The SDC applies special procedures to manage data with disclosure risk, such as carefully vetting researchers, requiring the signing of an agreement and applying various measures of organisational and technical control to protect data. The service is not yet fully set up and under development, so while the first users had already been granted access to the data, further steps are now being taken to improve procedures for secure data handling, better training of staff with regards to secure data handling and the consequences if users do not comply with the agreed procedures.

The user contract for the use of confidential data contains different measures in case of non-compliance: In case of misuse the user has to delete all data and supplementary material. In addition, a report will be sent to other data service centers as well as to the German Data Forum (RatSWD). It is also possible that the data user's account is blocked temporarily or permanently.

Links:

Data Catalogue and other portals: <http://www.gesis.org/en/services/research/>

Registration: <https://dbk.gesis.org/register/>

Terms of use/ usage regulations:

<http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/>



Charge regulations: <http://www.gesis.org/en/services/data-analysis/data-archive-service/charges/>

Contract for on-site use:

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_SDC\\_Data\\_Use\\_Agreement\\_On\\_Site.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_SDC_Data_Use_Agreement_On_Site.pdf)

Contract for off-site use:

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_SDC\\_Data\\_Use\\_Agreement\\_Off\\_Site.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_SDC_Data_Use_Agreement_Off_Site.pdf)

Technical and organizational measures (in German):

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_Hinweise\\_zum\\_SDC\\_Gastarbeitsplatz.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_Hinweise_zum_SDC_Gastarbeitsplatz.pdf)

Security requirements for off-site access (in German):

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure\\_data\\_center/GESIS\\_Merkblatt\\_Umgang\\_Forschungsdaten.pdf](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/secure_data_center/GESIS_Merkblatt_Umgang_Forschungsdaten.pdf)

All URLs accessed 07 April 2014

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

**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:*

All data users must agree to the usage regulations, which determine who can access data under which conditions. The Data Archive provides guidance on how to use and cite data obtained through GESIS. These guidelines are available on the website.

Individual-level data in social science research concerns behaviour, opinions, attitudes, as well as social and economic living conditions of individuals. These contents are all subject to codes of conduct for the protection of the participants in research projects. Thus, personality rights, e.g. the right to informational self-determination or the right to privacy, have to be considered.

Individual-level data is especially protected by the Federal Data Protection Act: “The purpose of this Act is to protect the individual against his/her right to privacy being impaired through the handling of his/her personal data” (BDSG §1, 1). For archiving purposes, the legal framework under which the data was collected must allow for data archiving and the individual’s right to privacy has to be protected. Usually data has to be anonymised.

The work of GESIS is based on a code of conduct and is explicitly bound to the rules of good scientific practice. An ombudsperson is in place who can be involved in the case of discrepancies concerning codes of conduct.

Studies deposited in the archive which contain confidential data are either anonymised or access is only given through our Secure Data Center, which provides controlled and secure access to data deserving special protection. Data depositors and producers are advised on dealing with disclosive data and anonymisation. Trained staff is available who can give advice on handling and processing of sensitive data as well as on accessing it through the Secure Data Center.

Checks for disclosure risks take place during the ingest phase and if problems are detected, data producers will be informed and advised. Anonymisation measures are carried out by archive staff as well.

Links:

Terms of use/ usage regulations:

<http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/>

Bibliographic citation of research data and study related documents:

<http://www.gesis.org/en/services/data-analysis/data-archive-service/citation-of-research-data/>

Regulations for citation (in German):

[http://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/\\_bgordnung\\_bestellen/Bibliographisches\\_Zitieren\\_von\\_Forsch](http://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/_bgordnung_bestellen/Bibliographisches_Zitieren_von_Forsch)

Guidelines of good scientific practice (in German):

[http://www.gesis.org/fileadmin/upload/institut/leitbild/Gute\\_Praxis\\_GESIS-Regeln.pdf](http://www.gesis.org/fileadmin/upload/institut/leitbild/Gute_Praxis_GESIS-Regeln.pdf)

Ombudsperson: <http://www.gesis.org/en/institute/the-association/>

Watteler, Oliver (2010): Datenschutz und die Archivierung von Daten in der qualitativen empirischen Sozialforschung. In: Medjedovi?, Irena; Witzel, Andreas; Mauer, Reiner (Mitarb.); Watteler, Oliver (Hrsg.): Wiederverwendung qualitativer Daten: Archivierung und Sekundärnutzung qualitativer Interviewtranskripte, Wiesbaden: VS Verl. für Sozialwiss., S. 55-94:

<http://www.springer.com/springer+vs/soziologie/book/978-3-531-15571-5>

**Data Seal of Approval Board**

W [www.datasealofapproval.org](http://www.datasealofapproval.org)

E [info@datasealofapproval.org](mailto:info@datasealofapproval.org)

All URLs accessed 07 April 2014

## **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 Archive strongly promotes data sharing and re-use and hence seeks to make data available as openly and as easily accessible as possible. However, legal regulations and respect for the needs and requirements of data depositors both make it necessary to manage access according to the principles outlined in our usage regulations and on the basis of an archive agreement.

If not indicated differently, the Data Archive makes data and documents available for scientific analysis carried out in academic research and teaching according to the stated access categories. Institutes and individuals outside academic research and teaching can apply for provision in written form.

Users have to register before they get access to the data. Upon registration they must accept the Archive's usage regulations.

The use of metadata from GESIS DBK is possible under a Creative Commons licence. All metadata are available free of restriction under the Creative Commons CC0 1.0 Universal Public Domain Dedication. However, GESIS requests that you actively acknowledge and give attribution to all metadata sources, such as the data providers and any data aggregators, including GESIS.

Links:

Terms of use/ usage regulations:

<http://www.gesis.org/en/services/data-analysis/data-archive-service/usage-regulations/>

Charge regulations: <http://www.gesis.org/en/services/data-analysis/data-archive-service/charges/>

Usage Guidelines for GESIS DBK Metadata: <https://dbk.gesis.org/dbksearch/guidelines.asp?db=e>

All URLs accessed 07 April 2014

## **Reviewer Entry**

*Accept or send back to applicant for modification:*

Accept

*Comments:*