Rules for Research Data

Decision-maker	Vice-Chancellor
Responsible unit	Research and Innovation Services
Decision date	2025-02-13
Period of validity	To be reviewed by 2028 at the latest
Summary	This document is an English translation of the rules on research data decided by the Vice-Chancellor. In the event of ambiguity, the Swedish version always applies. The document describes the University of Gothenburg's rules for the management of research data. The rules apply to planning and documenting data management as well as to the storage, collection, analysis, publication, and archiving of research data.

Introduction and Background

The University of Gothenburg's rules for research data aim to ensure that research data is managed responsibly and made accessible for use and sharing in accordance with the FAIR principles.[1]

An important purpose of these rules is to provide support for collecting, managing, storing, and sharing research data, thereby promoting openness, integrity, and quality in research. By adhering to these rules, University employees ensure that research data is preserved, organized, and made available for new research purposes in a well-documented process. The rules for research data apply to all aspects of research data management at the University of Gothenburg and are based on three central starting points:

- The University shall, based on legal requirements and these rules, provide the necessary infrastructure for storing, analysing, processing, and sharing research data during and after the completion of research activities,
- Information classification is a fundamental prerequisite for managing research data in a secure and appropriate manner,
- Research data should normally be made available through formal data publication.

The management of research data is also governed by legislation and other policy documents at the University of Gothenburg, for instance concerning information security, research ethics, IT resources, and more.

Data Management Plan

At the University of Gothenburg, data management plans shall be established for all research projects and other research activities. A data management plan (DMP) is a tool for documenting how research data is managed before, during, and after the research process. It provides a structured approach to documenting legal and regulatory compliance as well as clarifying the project's research data management. By drafting, maintaining, and utilizing a data management plan, good data management practices are promoted. The data management plan must be registered in the University's records (diarieförd) and can be updated as necessary. The principal investigator (PI) is responsible for creating, maintaining, and registering the plan.

Storage of Research Data

Research data in digital formats must be stored using the University's designated storage services or specific IT environments provided by the University for individual projects. In some cases, it may be warranted to store research data externally—for example, in established international research databases or large-scale storage services connected to research infrastructures or research centers. In such cases, it is important to ensure that agreements are in place to regulate personal data processing, data sharing, archiving authority, and that other legal and regulatory requirements are met.

Research data in non-digital formats must also be stored, managed, and preserved in a way that meets the requirements for the material's information classification. The principal investigator is responsible for documenting the information classification and ensuring that storage takes place in one of the University's designated storage services, and that the chosen storage service is approved for the relevant information classification.

[1] https://www.go-fair.org/fair-principles/

The University is responsible for providing general storage services as well as administrative support and technical support.

Collection and Analysis of Research Data

In the collection, analysis, and processing of research data, tools adapted to the material's information classification must be used. Primarily, the tools provided by the University should be utilised. In cases where personal data will be processed, the University's guidelines for the processing of personal data for research purposes must be followed.

Publication of Research Data

Research data shall be published as independent data publications in order to promote open science and reuse. Data publication is not the same as the publication of research results or scientific articles. Data publications make it possible to cite published data and to link data to researchers and research projects.

Where possible, research data should be made openly available upon publication. Data that cannot be made openly available should be published in the form of a metadata publication referencing restricted data files that are prepared for reuse. Research data are public documents ("allmänna handlingar"), whether or not they are formally published; requests for their release are assessed according to the Swedish Public Access to Information and Secrecy Act and are handled in line with the University of Gothenburg's routines.

Rich metadata and detailed documentation should accompany the data publication. The data should be handled in such a way that it remains reusable without technical barriers in the future. Published data and metadata should be assigned persistent identifiers (PIDs). Persistent identifiers referring to individuals, organizations, and publications are particularly important to include. Wherever possible, metadata should be machine-readable. In most cases, metadata and documentation should be fully disclosed even if the data is restricted. Software associated with the data publication should, if possible, be published as open-source code along with reproduction examples.

A suitable platform for data publication should be chosen based on the nature of the dataset and any subject-specific needs. As a first option, platforms that provide the ability for rich metadata annotation and that apply an independent, manual quality review process before publication should be used. A recommended Swedish data repository is provided by the Swedish National Data Service (Svensk nationell datatjänst, SND).

Archiving of Research Data

Research data must be preserved in accordance with the Swedish Archives Act (Arkivlagen) and the University's decisions regarding the application of the Swedish National Archives (Riksarkivet) regulations. The general rule is that research data shall be preserved indefinitely unless a decision for disposal ("gallring") is made. Such a decision can only be made after—at the earliest—ten years, depending on the type of data and the type of project funding. To determine whether research data should be preserved or may be disposed of, an assessment (bevarande- och gallringsutredning) must be conducted, based on the criteria established by the Swedish National Archives, which evaluates the ongoing value of the research data. Research data that may have continued scientific value or may be of historical or public interest must be preserved. Examples of such documents include:

- Particularly extensive primary data that is unique or that can only be recreated with great effort.
- Registers and databases containing data with especially high coverage and verifiability.
- Documents that shed light on the historical development of a scientific discipline.
- Documents that illuminate the academic environment from a cultural-historical perspective.
- Documents that highlight the activities of a notable individual.
- Documents concerning research that has received significant attention in public debate or is expected to do so once its research results reach wider dissemination.

A preservation and disposal assessment is conducted in consultation with the University's archives function. If the assessment concludes that the research data can be disposed of, a disposal decision must also be made by the responsible Head of Department (prefekt) or equivalent.

Appendix:
Definitions and Explanations of Terms

Research Data

Research data includes all materials and information, regardless of format, that have been collected or generated for scientific analysis. This can range from interview recordings, images, and films to measurements, experimental results, physical documents, observations, or survey data. Research data may originate from quantitative as well as qualitative processes, and may consist of primary data gathered specifically for a research project or secondary data reused from previous studies.

Information Classification

Information classification is the process of categorizing information based on its sensitivity and its value to an organization, which determines how it should be protected. Through classification, the necessary security measures are identified to protect the information from unauthorized access or disclosure. The University of Gothenburg's information classification model is used for this purpose. The individual responsible for conducting the information classification is the information owner.

Public Document ("Allmän handling")

A public document is any document or recording (for example, an audio or video recording) held by a public authority and which has been received or drawn up by that authority. A recording is in the authority's possession if it is accessible to the authority via technical means that the authority itself uses to convert the recording into a readable, audible, or otherwise perceivable format. Public documents, such as research data, may be requested through a formal inquiry to the University, in accordance with the Swedish Freedom of the Press Act (Tryckfrihetsförordningen). Each disclosure request is assessed under the Public Access to Information and Secrecy Act (Offentlighets- och sekretesslagen) and handled according to the University of Gothenburg's procedures for releasing public documents.

Open Access to Research Data and the FAIR Principles

The Swedish government has declared that, as far as possible, research shall adhere to the FAIR principles. This means that research data should be published "as openly as possible and as restricted as necessary." FAIR stands for Findable, Accessible, Interoperable, and Reusable. Publishing in line with the FAIR principles does not necessarily mean that data must be openly available. The goal is to ensure that research data are well-documented and to enable effective reuse.

Research Projects and Research Activities

In these rules, research projects and research activities are understood as defined, delimited activities carried out at the University by a researcher employed at the University (or under the leadership of such a researcher), with the purpose of producing publishable scientific results (or similar outputs).

DMP

An abbreviation of the English term "Data Management Plan," sometimes used in Swedish synonymously with "datahanteringsplan."

The University's Guidelines for the Processing of Personal Data for Research Purposes

The University's legal team has prepared a series of checklists for processing personal data that clarify and simplify this work. These are published on the University's staff portal. The checklist for information owners within research, as well as the checklist for system owners (with regard to new digital tools used in research), contains a number of steps for anyone responsible for the processing of personal data. The checklist is based partly on the legal requirements set out in, among other instrument, the General Data Protection Regulation (GDPR), the Public Access to Information and Secrecy Act (OSL), and the Swedish Civil Contingencies Agency's regulations (MSB), and is partly intended to protect the University's valuable assets.

Publication

In this document, the term "publication" mainly refers to the publication of research data, whether openly accessible or as a metadata publication. Data publication is not synonymous with the publication of research results and scientific articles. Data publication is typically done in a data repository.

Rich Metadata

Examples of rich metadata and thorough documentation include markup using controlled vocabularies, explanations of the content of each variable, summary documentation to enable reuse of the data, and a description of how the dataset in question can be accessed.

Persistent Identifiers

A technology for unambiguously designating a resource or phenomenon, now and in the future. The identifier can, for example, specify a publication, person, or organization along with the resource's metadata. Examples of persistent identifiers include DOI, ORCID, or ROR ID.

Reproduction Examples

A reproduction example usually contains source code, documentation, references to or copies of input, and examples of output.