Research Data Management and Open Science Policy Faculty of Geosciences

Geo Data Team (GDT)/Faculty of Geosciences/Utrecht University

Version: 1.0

March 2025

Somewhere, something incredible is waiting to be discovered—but only if we make data accessible, reusable, and open to all.

— Adapted from Carl Sagan

Contents

Version History and Document Revisions

	Version	Changes	Date
1	**V0.0**	The initial version	March 2023
2	**V0.1**	The feedback from the following departments have been implemented: • Geo-ICT (Peter Hessels) • Faculty policy advisor (Eveline Helsper) The document's structure has been modified and divided into three sections: Introduction, Policy Statements, and Procedures.	June 2023

3	**V0.2**	The feedback from the following units have been implemented:
		• Research Support Office (RSO)
		  ○ Oscar van Vliet
		• Research Coordinators
		  ○ Marina Jongkind
		  ○ Eldert Advokaat
		The following changes have been made:
		• Abstract is added
		• The entire Introduction chapter has been revised
		• Section 2.2 has been revised
		• Section 3.1.2 has been revised
			November 2023

4	**V0.3**	The abstract, section 1.1, and the contact information for reporting the data breach have been revised based on the feedback provided by: • Peter Hessels (Head of Geo-ICT) • Vincent Brunst (Data Coordinator)	January 2024
5	**V0.4**	The feedback from the department board members has been incorporated. This feedback was gathered during the policy presentation to the following departments: • Sustainable Development (2024-02-13) • Human Geography and Planning (2024-03-05) • Earth Sciences (2024-03-11) • Physical Geography (2024-05-22)	November 2024
6	**V1.0**	The feedback from the faculty board members has been incorporated. The feedback was provided by:	March 2025

		• Marc Bierkens (Interim dean of the faculty) Based on his suggestion, two statements about the misuse of data by data publishers and data users have been added to the Ethics and Data Privacy section. The feedback from two research staff of two departments has been incorporated: • Emilia Jarochowska (Department of Earth Sciences) • Marco Helbich (Department Social Geography; SGPL)	

Acronyms & Glossary

- AVG: Algemene Verordening Gegevensbescherming (Dutch name of GDPR)
- UU: Utrecht University
- CIA: Confidentiality, Integrity, Availability
- DAG: Data Archive Geosciences
- DMP: Data Management Plan
- DOI: Digital Object Identifier
- ERB: Ethics Review Board
- FAIR: Findability, Accessibility, Interoperability, Re-usability
- GDPR: General Data Protection Regulation
- GDT: Geo Data Team
- HE: Horizon Europe
- HR: Human Resources
- ICT: Information and Communications Technology
- MFA: Multi-Factor Authentication
- PI: Principal Investigator
- RDM: Research Data Management
- RSO: Research Support Office
- VPN: Virtual Private Network
- WBS: Work Breakdown Structure (Budgeting terminology)

- **Data Management Plan (DMP):** A document outlining the procedures for handling data throughout a research project and post-project completion. Typically, it comprises the following sections: data collection, storage, and backup; documentation; access; sharing and reuse; and preservation and archiving.
- **Open science:** Practicing science in a sustainable manner that allows others to work with, contribute to, and use the scientific process.
- **FAIR principles:** The FAIR principles refer to a set of guiding principles for scientific data management and stewardship. The acronym FAIR stands for Findable, Accessible, Interoperable, and Reusable. These principles aim to make research data more discoverable, understandable, and usable, which can enhance the efficiency and effectiveness of scientific research and aim to promote good data management practices that support the scientific enterprise, enable data sharing and reuse, and ensure that data is used ethically and responsibly.
- **Personal data:** Any information that relates to an identified or identifiable natural person. It can include any information that is linked or related to a person either directly or indirectly.

- **Privacy scan:** A tool developed at the Faculty of Geosciences designed to facilitate the process of evaluating and documenting a project's GDPR compliance.
- **Data breach:** A data breach is an unauthorized access, disclosure, or loss of sensitive information, often involving personal data which can result from cyberattacks, hacking, human error, or physical theft of devices containing sensitive data.

Chapter 1:

Introduction and Context

1.1 Abstract

The current Research Data Management (RDM) and Open Science activities at the Faculty of Geosciences involve a holistic approach to data management, Open Science, and privacy. However, there is a missing framework addressing critical components—such as proper research data handling, responsibilities of different stakeholders, and training and support—and a structure to ensure a comprehensive and effective RDM and Open Science strategy at the faculty. Furthermore, the recognition and rewards system in academia is evolving towards broader metrics, including Open Science practices, recognition of interdisciplinary and collaborative research, and a focus on the societal impact of research. A well-defined Open Science and Research Data Management (RDM) policy positions our institute to navigate these changes effectively, ensuring that our research activities are not only aligned with emerging criteria but also actively contribute to the advancement of science in innovative and impactful ways.

This policy provides the necessary framework to assist and empower researchers in achieving compliance with Open Science and RDM principles, drawing inspiration from both the Open Science and RDM requirements stipulated by grant funders and the overarching high-level policies established at the national level. At its core, this policy emphasizes the roles and accountabilities of all stakeholders, from individual researchers and research support staff to the faculty as a whole. It underscores the transformative potential of research data management guided by the FAIR principles (Findable, Accessible, Interoperable, and Reusable) and Open Science standards. By fostering robust data handling practices from project inception to dissemination, the policy seeks to maximize the visibility, reproducibility, and impact of research outputs. Moreover, it outlines actionable steps to embed these principles into everyday practices, leveraging the resources and facilities available at Utrecht University and the Faculty of Geosciences. Through this framework, the policy equips the research community to meet the expectations of grant funders and other key stakeholders while also strengthening collaboration, innovation, and the overall quality of research.

To ensure the successful implementation of the policy and empower all those involved, structured support will be provided through comprehensive training programs, access to tailored resources, and expert guidance. These activities will be integrated into a detailed implementation strategy and communicated clearly to all stakeholders.

1.2 Purpose

Although the Faculty of Geosciences has made significant progress in the last 5 years in promoting Research Data Management (RDM), Open Science, and data privacy, a structured framework is still needed to effectively integrate and address these critical components. Recognizing this need, we have begun developing a robust framework designed to integrate essential elements and structures, strengthening the foundation for successful and impactful

RDM and Open Science practices within the Faculty. Furthermore, the recognition and rewards system in academia is evolving, with a potential decrease in the exclusive reliance on metrics like the H-index and the number of publications. Emerging trends suggest a move towards broader metrics, including Open Science practices, recognition of interdisciplinary and collaborative research, focus on the societal impact of research, etc. (See The Leiden Manifesto for research metrics, The Hong Kong Principles for assessing researchers, and Dutch Recognition & Rewards programme). A clear Open Science and RDM policy equips our faculty to adapt to evolving standards while empowering our researchers to drive innovation and make meaningful contributions to the advancement of science.

This policy establishes a framework of principles, supported by detailed guidelines and procedures, to ensure effective research data management within the Faculty of Geosciences. It aims to assist researchers, staff, and affiliated individuals in managing research data throughout its lifecycle—from creation and collection to analysis, preservation, and sharing.

The policy is designed to:

- Ensure compliance with FAIR and Open Science principles, grant funder's requirements, and applicable data protection and ethical standards.
- Promote data sharing and collaboration within the institution and with external partners.
- Clarify roles and responsibilities for all stakeholders in implementing and adhering to the policy.

To support its implementation, the policy provides access to training, resources, and ongoing support, creating a foundation for consistent, ethical, and impactful research data practices.

1.3 Scope

This policy applies to all research data, including but not limited to raw data, metadata, laboratory records, field observations, models, simulations, and codes, generated or collected by individuals affiliated with the Faculty of Geosciences at Utrecht University during their research activities. This encompasses data sourced from various research methods, including laboratory experiments, fieldwork, surveys, and simulations. The policy applies to all stages of the research data lifecycle, including data collection, storage, analysis, sharing, and long-term preservation. It aligns with the Strategic Plan 2025, the policy framework of research data, Dutch code of conduct for scientific integrity, and the Open Science ambition of Utrecht University. It applies to both internally generated and externally collaboratively produced research data, addressing the ownership, access rights, and outlines the expected data preservation period in compliance with relevant legal and regulatory requirements, including GDPR, data protection laws, intellectual property regulations, and ethical guidelines.

1.4 Target Audiences

This policy has diverse audiences owing to its interdepartmental scope. These are Researchers and all individuals involved in the research process, including but not limited to educators, students, lab managers, technicians associated with the Faculty of Geosciences, the Research Support Office (RSO), Geo-ICT, Geo Data Team (GDT), and the Faculty's Management at different levels.

1.5 Overview of the Policy Document

The policy consists of four chapters: Chapter 1 will discuss the purpose and scope of the policy of the Faculty of Geosciences. Chapter 2 introduces and elaborates upon the policy statements for different topics in RDM, Open Science, and data privacy. Chapter 3 outlines detailed procedures and actions to operationalize the policy statements. Chapter 4 contains extra information and appendices related to the policy.

Chapter 2:

Policy Statements

2.1 Data Ownership and Responsibilities

- Research data generated or collected by employees, researchers, and affiliated individuals during their association with the Faculty of Geosciences is, in principle, owned by Utrecht University. For more information, please refer to the Dealing with intellectual property rights page on the intranet.
- The principal investigators, project leaders, supervisors, Postdocs, and Ph.D. candidates are responsible for ensuring that research data generated or collected by them or other individuals under their supervision—including Master's and Bachelor's students— adhere to this and other relevant policies at Utrecht University, as well as relevant legal, ethical, and funding requirements.
- The research data should be properly stored in a trusted and secure storage solution to ensure long-term accessibility even if the primary researcher leaves the Faculty of Geosciences. The data should not be stored solely on personal accounts or devices following the research data management best practices outlined in this policy.

2.2 Data Management Planning

- Researchers are required to create a data management plan (DMP) for each research project, whether it is externally or internally funded based on the template provided by the grant funder, Utrecht University, or other officially accepted templates.
- The DMP must be submitted for review and approval to the GDT before starting data collection.

2.3 Data Storage and Security

- Research data should be stored in secure and backed-up systems provided or approved by Utrecht University or the Faculty of Geosciences.
- The storage solution should be selected based on the size, data type, cost, sensitivity, processing, and data access activities.
- Data must, whenever possible, be stored in file formats that are open, non-proprietary, and easily accessible, to facilitate sharing, reusing, and long-term preservation.
- Data storage might have some associated costs, and it is recommended to include the data management costs in the budget table of the relevant research proposal.
- Adequate security measures, including access controls and encryption, must be implemented to protect research data from unauthorized access, loss, or alteration.

2.4 Data Documentation and Metadata

- Researchers must document research data and create associated metadata, including details about data collection methods, instruments, variables, and data processing procedures.

- When data is shared or publicly released, accompanying documentation must be provided to ensure proper interpretation and enable appropriate re-use of the data for future purposes.
- The provided metadata should be human- and machine-readable in the common formats of the research field or data types to facilitate data discovery, understanding, and reuse.

2.5 Data Sharing and Access

- Researchers are encouraged to share research data with internal and external colleagues as soon as possible, subject to legal, ethical, and privacy considerations.
- A data sharing agreement should be established when sharing data with external parties, especially those outside the European Union.
- Researchers involved in research projects should share research data with project members as needed, ensuring compliance with data privacy and security guidelines.

2.6 Data Publication and Archiving

- Research data should be published as openly as possible in accordance with the principles of Open Science and FAIR and subject to legal, ethical, and privacy considerations.
- Research data should be published on trusted data repositories or platforms, which provide DOIs and have a peer-review process.
- Research data that is considered necessary to safeguard the integrity of the research project findings should be retained for a period specified by funding agencies, legal requirements, or UU policies.
- Research data should be published with an appropriate license to specify the conditions for reuse.
- When research data cannot be published openly for reproduction and re-use purposes due to legitimate reasons, it should be deposited in a data archive facility following the FAIR principles.

2.7 Data Privacy and Ethics

- All the research projects that process personal data must have appropriate documentation for compliance with the GDPR and must follow data protection legislation of other countries when relevant.
- Appropriate data privacy compliance documentation must be in place before the start of the project, and it must be reviewed and approved by the faculty privacy officer.
- Everyone involved in the processing of personal data must receive sufficient training in handling such data.
- All research projects with potential ethical issues should obtain ethics approval. For research related to human subjects, this can be requested from the Science-Geo Ethics Review Board (SG-ERB) before the project's commencement.

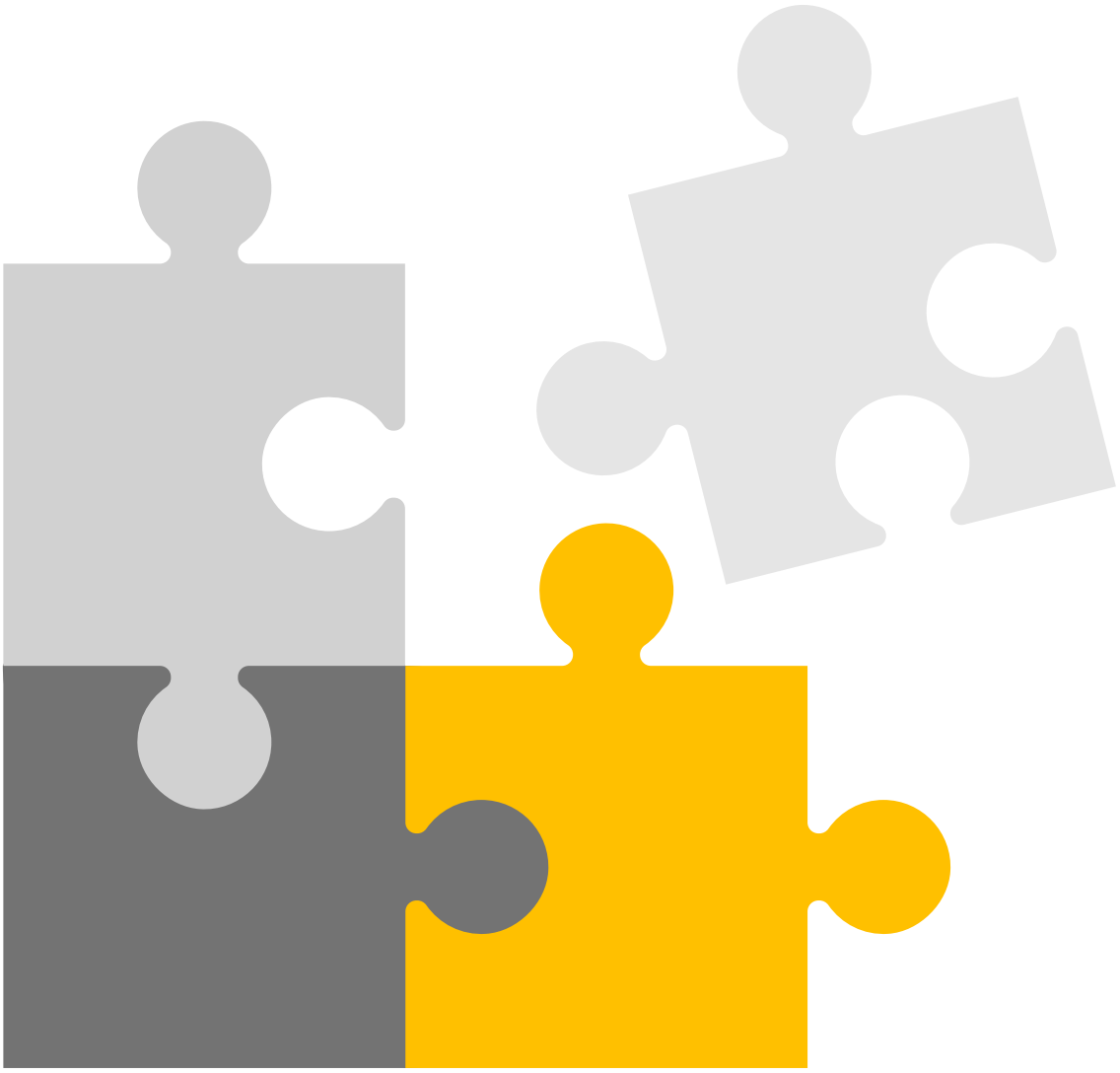
- Researchers must prevent misuse and dual use of data by assessing risks and applying safeguards. When reusing others' data, they must ensure proper attribution, respect licensing terms, and comply with ethical and legal guidelines.

2.8 Training and Support

- GDT will provide training and support to researchers on research data management best practices, Open Science, data protection compliance, and relevant tools and resources.
- If research projects require dedicated regular support, GDT can provide it in terms of resource availability.

Chapter 3:

Procedures



3.1 Data Management in Grant Proposals

3.1.1 Data management section in proposals

The grant proposal's data management section needs to be filled out to demonstrate how the research data will be managed and preserved during the research project. This part of the proposals should be written based on the latest national and European RDM policies. The following items must be considered:

- Researchers preparing a research proposal should contact the GDT in a timely manner to draft or review the data management section.
- The RSO funding advisor will address the need for support in writing the data management section, ensuring that GDT is involved in a timely manner (preferably at least 2 weeks before the submission deadline).

The time required to review the data management section for each grant call may vary. For most of the national calls, GDT processes the requests and provides feedback on the proposal within 5 working days.

Note: Currently, some grant calls, e.g., Horizon Europe (HE), evaluate the data management, open science, data security, privacy, and ethics sections of the proposals in the early phases of reviewing. For these calls, it is highly recommended that the applicant starts to set up the mentioned sections as soon as possible since the time needed for GDT to review will be longer.

3.1.2 Data Management Plan (DMP)

Drafting a well-structured and comprehensive Data Management Plan (DMP) serves as the initial and foundational step in establishing effective data management practices and plays a pivotal role in ensuring data integrity, security, and accessibility throughout the research project's lifecycle. It serves as a dynamic roadmap for researchers, project teams, and relevant stakeholders, offering essential insights into how data will be handled, thus promoting the long-term utility and reproducibility of research findings. Below, you'll find a step-by-step guide on how to create and manage DMPs at the Faculty of Geosciences:

- The project coordinator of a research project (in most cases, the principal investigator or project manager) is responsible for preparing a draft of the data management section and data management plan (DMP) using available manuals and in consultation with GDT.
- For most of the grant calls, it is required to submit the first version of the DMP up to 6 months after starting the project.
- At Utrecht University, all DMPs should be made via the DMPOnline platform. Guides for using DMP Online are available on the GDT website. The DMP templates of the major funders, including NWO, ERC, and Horizon Europe, are available on DMPOnline.
- The project coordinator makes sure that a DMP will be set up and shared with the data stewards of GDT. The designated member of the GDT reviews the DMPs and gives feedback at least 2 months before the deadline.

- GDT is responsible for reviewing the DMP during the project for compliance with funder requirements, FAIR, Open Science, data protection principles, and other relevant guidelines.
- If a Data Management Plan (DMP) does not specify requirements for data versioning, folder structure, and documentation, researchers must still ensure that their data is well-organized, traceable, and understandable to facilitate collaboration, reproducibility, and long-term accessibility.

3.2 Data Storage and Security

3.2.1 Data Storage

- Storing your research data in a way that ensures safe backup is a key step in effective data management. The selection of the most appropriate storage solutions for your research data depends on several factors, such as the data type, the size, the performance, the accessibility, etc. For your convenience, you can use the Data Storage Finder or see Table 1 as a selection guide for the proper data-storing solution:
- **Small data (< 50 GB):** In general, small datasets can be stored on OneDrive, TeamSite/SharePoint, Yoda, and SURF Drive.
- **Large data (> 50GB & < 1 TB):** Large datasets can be stored on OneDrive, TeamSite/SharePoint, SURF Drive, and Yoda.
- **Big data (>1 TB):** Big datasets should be stored on Yoda and UU's institutional research drive.

To find a proper data storage solution for research data, the guidelines below should be followed:

- To choose the most suitable storage solution for each project, a designated member of the GDT will consult the project representative upon request.
- To use Yoda, the PI or the main researcher sends a request to datateam.geo@uu.nl with the required information, and access will be granted within 1-2 working days.
- When support is needed for arranging a (different) storage location for already produced research data, GDT will provide advice.
- The involved researchers under the supervision of the PI are responsible for the content of the data, storing the data in the (agreed) folder structure, and applying (agreed) conventions in folder and file naming and proper documentation to interpret the stored data correctly.
- Every storage location has a data owner, who is responsible and determines which individuals can access the data. In addition, every storage location has an expiry date, so it is clear when data should be transferred to solutions prepared for long-term preservation.
- GDT can advise and support in selecting or designing an appropriate storage location (related to other services, like computing power, web hosting, etc.), setting up folder

structures, naming conventions, and preparing data documentation. It is up to the researcher involved if advice or support given will be implemented in the project.

- When hardware must be purchased to set up a storage solution, a request must be made to the faculty ICT demand manager. He will retrieve a quote and discuss the details with the researcher involved before any order is given to the designated supplier.

3.2.1.1 Where not to store data

Data should not be stored solely locally on computers, USB drives, or external hard drives, these devices are unreliable and can be easily lost. In case it is necessary to store personal, sensitive and / or valuable data on one of these devices, the device must be fully encrypted using operating system-level encryption such as Windows BitLocker. Some USB drives come with encryption; however, these encryption schemes do not usually meet a good level of data security and protection.

Data owned by UU should also not be stored on cloud services for which UU does not have an official agreement or contract. This includes but is not limited to Google Drive, Dropbox, Mega, and Box.com. If another institution is exclusively handling a project's data management, such as in a consortium, then the use of other tools agreed with and approved by the consortium partners may be permitted.

3.2.1.2 Storing Informed Consent Forms

When you have obtained either filled informed consent forms or established consent for obtaining data during the collection process, explicitly mention where they will be stored. This can be done in a research protocol, DMP, and / or privacy scan.

Make sure that the consent data will not be stored together with the research data and that it is only accessible to the project members on a need-to-know basis. When the consent data contains personal data, make sure that the corresponding files are encrypted.

3.2.1.3 Costs of data storage

There are different storage solutions available at Utrecht University. The costs of using each of them are organized differently and the user must check with the Geo Data Team before deciding which storage solution will be used:

- The costs of using UU's recommended solutions, like OneDrive, Teams / SharePoint, SURF Drive, SURF Research Drive, and U-drive are covered by the UU. These solutions have a maximum storage capacity that initially cannot be increased.
- The costs of using internal data storage solutions like O-drive and Yoda, are now covered by the faculty. There are agreements to charge the costs internally to the departments, based on the current use. When substantial amounts of data (multiple TBs) for a single project or program for a longer period need to be stored on these solutions, specific arrangements need to be set up. This may also include charges to the budget number (WBS) of a research project. To determine if there are costs

involved, that should be covered by project or individual budgets, it is strongly advised to contact GDT beforehand.

- It is highly recommended to include the data management costs in the grant proposal's budget table to ensure that the costs of data storage, archiving, and other data-related costs will be covered. GDT provides support in estimating data management costs in grant proposals.

3.2.2 Data Security

Data security is an essential part of data management, especially in today's digital age, where data breaches and cyber-attacks are prevalent. The following are some key points recommended to be followed at the Faculty of Geosciences:

- Research data should be transferred from portable devices to trusted UU data storage platforms as soon as possible following the guidelines for handling data in portable hard drives.
- Data should not be stored locally on computers, USB drives, or external hard drives, as these devices are unreliable and can be easily lost. If data is stored on these devices, the device should be fully encrypted using operating system-level encryption.
- For the data storage platforms, whenever possible, use multi-factor authentication (MFA, sometimes also known as two-factor authentication or 2FA). For UU services that require MFA, you should activate it for your UU account. Also, external users who are added by UU staff to Microsoft services (e.g. Teams, SharePoint, OneDrive), should have activated MFA linked to their Office365 account.
- A proper backup strategy should be implemented in all platforms used for storing research data. To back up research data effectively, it is strongly recommended to adhere to the 3-2-1 rule, which dictates having: 3 copies of your data, utilizing 2 different mediums (such as distinct storage types or models of hard drives), and keeping 1 copy off-site to guard against natural disasters. If your chosen data storage solution does not align with the 3-2-1 rule, it is essential to plan for backup independently.
- Only trusted applications should be used for data processing/sharing and tools that might cause any data breach should be avoided. Each dataset should be classified according to the CIA triad (Confidentiality, Integrity, Availability). Procedures to store and get access to data sets will be set according to the CIA classification.
- Classifying your data and knowing which corresponding security measures to take is something you do before each new data processing activity. More information about data classification can be found here.
- In the event of a computer security incident, promptly report it to cert@uu.nl or, in urgent situations, call +31 (30) 253 5959. Such incidents include, but are not restricted to, password breaches, lost computers and phones, as well as theft or damage to IT

assets. If you come across phishing attempts or suspicious emails, report them to phishing@uu.nl.

- If you suspect a data breach at Utrecht University, whether it arises from unintended incidents involving personal data, unauthorized access, data loss, or the misplacement of UU devices, it is crucial to report it immediately. You can do so by contacting the ICT Service Desk via email at datalek@uu.nl or, in emergency cases, by calling +31 (30) 253 4500.
- Do not share your passwords and use passwords that are at least 12 characters in length consisting of letters, numbers, and symbols.
- Check to make sure your computing devices use whole drive encryption to protect the data from unauthorized access if a device is lost or stolen. Lock your device when you leave your workspace, even when your co-workers are around.
- Utrecht University is ultimately responsible for safeguarding personal data, and as an institution, it relies on researchers to keep data safe, and in compliance with the rules of the GDPR. Consult the GDT to find the appropriate locations to store and process data, use only approved tools for processing data, and inform the data protection officer as soon as a data breach is known.
- The GDT and Geo-ICT present data security best practices to the faculty members, especially for new employees, regularly via different communication channels.

3.3 Data Documentation and Metadata

Documenting data is one of the basic steps for managing data properly and following the FAIR principles, especially in increasing the reusability of the data. At the Faculty of Geosciences, it is highly recommended to include a README file or any type of documentation with the data or any research output. Also, providing a complete set of metadata together with data helps others to find the data set and understand what it is about. GDT facilitates creating metadata and documentation by providing training and consultancy to researchers on how to document data efficiently and professionally.

- The researcher and PI are responsible for providing documentation and rich metadata together with the data. How-to manuals are available via the Geo data support website, as well as workshops that the GDT organizes regularly.
- GDT reviews the quality of documentation and associated metadata and their compliance with FAIR principles at the data publication phase.

3.4 Data Sharing and Access

At Utrecht University, sharing data is valuable as it is one of the main pillars of Open Science and is ethically required. However, for sharing data, there are some guidelines and requirements that need to be considered in the data-sharing process:

- During the research project, data and related materials should be shared with at least one of the group members, e.g., the principal investigator, supervisor, or other

partners in the project, via one of the trusted data-sharing platforms at Utrecht University. The trusted data storage platforms with data-sharing capabilities at UU are listed in Table 1.

- Data should be documented properly before sharing, e.g., using README or any other type of associated metadata.
- All employees at the faculty must use trusted tools for sharing data, e.g., SharePoint, Yoda, SurfDrive, etc. Trusted tools can be found with the UU ICT Tool Advisor or by asking the GDT.
- Personal and sensitive information in research data must be shared in compliance with the GDPR (The European General Data Protection Regulation, AVG in Dutch).
- The corresponding person (e.g., PI or appointed person for data management) should check whether there is any personal data to share and make sure they are not violating the data privacy regulation before sharing any data.
- In the case of a collaboration with a third party in a research project, an article must be made for data sharing in the contract between the consortium and the involved third parties. The Geo Data Team provides support to draft a tailor-made agreement based on the UU standard template for data-sharing agreements.

3.5 Data Publication and Archiving

3.5.1 Data publication

Publishing research data plays a crucial role in advancing scientific knowledge, promoting transparency and accountability, facilitating collaboration, increasing resource optimization and sustainable research, and driving innovation across various fields. Also, publishing data is one of the requirements from the research grant funders and is as important as publishing articles. To publish data properly, there are guidelines that should be followed:

- Deposit and publish your data to a domain-specific repository, if available, which provides a Digital Object Identifier (DOI). If there is no established domain-specific repository, please use Yoda—UU's institutional repository— to publish your research data.
- The data publication workflow is different among data repositories, and the data might be checked in terms of quality. It is recommended to ask GDT to review datasets before publishing.
- In Yoda and DataverseNL, GDT or RDM support review datasets before publication.
- Besides the data itself, a data documentation file (e.g., README.txt) and rich metadata should be provided.
- Research data should be licensed at the time of publication under CC-BY or similar open-access licenses. If there are valid reasons to restrict the reuse of the provided data, use another suitable license. There are different data licenses, for support to choose which license is appropriate for your dataset, see the data license webpage.

- If your dataset is derived from another data publication, you must include the citation of this dataset in the accompanied data documentation and metadata.

3.5.2 Data archiving

To ensure the enduring preservation of research data beyond the duration of the project in which it was originally generated, the following steps should be taken:

- Utilize a storage platform that is well-suited for the long-term archiving of research data. The data from the storage systems listed in Table 1 which are better suited for the retention of active data, should be transferred to the long-term archival storage platform at the end of the project.
- Before the archival process, curate your data by identifying the data that is appropriate for public release, retention for reproducibility or reusability, and those that are no longer necessary.
- The researcher(s) and the PI will decide which data archiving platform will be used for the long-term preservation of the data. To facilitate the decision procedure, the Data Archive Geosciences (DAG) platform is promoted for data archiving by the Faculty, and GDT provides consultancy and organizes workshops related to this topic.
- Data that is no longer necessary should be deleted from the devoted storage locations to free up storage space for future research data.
- For the preservation of static research data that is not publicly released and is relevant for sharing among colleagues or other interested parties within the faculty for reproducibility and reusability purposes, it is recommended to use Data Archive Geosciences (DAG).
- Research data already stored in Yoda, can be retained there for long-term preservation. To ensure its proper management, it is recommended to fill in the metadata form. There is also an option to transfer your data to Yoda Vault, where it will be rendered as read-only to prevent any tampering.

3.6 Data Privacy and Ethics

3.6.1 Data privacy

All projects at the Faculty of Geosciences that process personal data must comply with the GDPR/AVG and must have written documentation of that compliance. It is advised to follow the guidelines below to ensure that the research project's activities are compliant with GDPR/AVG:

- The Geo Privacy Scan is the preferred tool to demonstrate data privacy compliance at the faculty. It is advised not to start collecting personal data if a Privacy Scan is not completed, to avoid the possible cancellation of the project.
- Everyone who is involved in the management and decision-making process of the project is responsible for learning how to handle personal data. The privacy workshops and the GDT website are good resources for learning about data privacy.

- The PI or responsible persons (project's controllers) are responsible for starting and maintaining a privacy scan document of the project, which documents GDPR compliance.
- A Privacy scan needs to be started at the same time the project is being designed. It must be continuously updated alongside the project life cycle until personal data has been deleted and/or fully anonymized.
- The faculty privacy officer, which is part of the GDT, is responsible for providing support to faculty members and for producing and maintaining training and awareness materials on data privacy.

3.6.2 Ethics

- All research projects processing personal data or involving human subjects must be reviewed by the Science-Geo Ethics Review Board (SG-ERB) before data collection starts.
- An ethics review of a research proposal can be requested by submitting an application form, together with other project documents, such as DMP, informed consent forms, information sheets, and Privacy Scan. More information about the working procedures and responsibilities of the SG-ERB can be found on the SG-Ethics Review Board web page.
- Obtaining ethics approval ensures proper assessment of risks before data collection or reuse. If you have any doubts about the ethical aspects and/or licensing conditions of reusing research data published by others, the GDT is your first point of contact for assistance.

3.7 Training and Support

GDT is the main point of providing support, consultancy, and training at the Faculty of Geosciences in research data management, Open Science, data privacy, and related topics.

- GDT is the first contact point for services related to research data management.
- GDT provides training and workshops regularly on different topics in research data management, Open Science, and data privacy for faculty researchers.
- Tailored pieces of advice on all the topics that are mentioned in this document are provided by GDT upon request.

3.7.1 Liaising data steward/manager with research projects

If a research project needs to have a dedicated data steward/manager to coordinate the data management activities, it is possible to designate one of the GDT members to collaborate part-time with the project. The details of such an allocation should be discussed between the project coordinator and the GDT. Consider that members of the GDT can only be designated in terms of availability. The request should be sent to the data coordinator, and a decision will be made in collaboration with the other stakeholders (e.g., GDT, Geo-ICT, HR).

Chapter 4: Appendices

4.1 Recommended Storage Solutions

Table 1: The recommended data storage solutions. The sharing column demonstrates with whom data can be shared and not the default sharing configuration.

Location	Max Storage	Suitable for sensitive data	GDT Recommended	Sharing	Description
OneDrive	1TB Shared	✓	✓	UU & Anyone external with an email address.	Both data storing and managing services are provided by Microsoft, for which a data processing agreement was negotiated with UU, that provides assurances regarding compliance with the GDPR and UU's privacy and security policies. You are the only one able to access these files, but access can also be granted to colleagues or external people. The two services are accessible within and outside UU through a web browser. Collaboration with others in UU or guests with an Office 365 account is possible.
SharePoint / TeamSite		⚠			TeamSite is an extension of SharePoint, and SharePoint is an enterprise data storage and group organization system. ⚠ Data in a SharePoint site is visible to anyone who is a named member of that SharePoint site.
SURF Drive	500GB	✓	✓	Editing: UU & Any SURF Participating Institution Read Only: Anyone	SURF Drive is a personal cloud storage service provided by SURF for every affiliated Dutch Institution or research community. You can access SURF Drive only with an affiliated Dutch Institution account. SURF Drive can be used to store, share, and synchronize data on the secure SURF cloud.
SURF Research Drive	Negotiable€	✓	✓	Editing: UU & Anyone associated with a UU project	SURF Research Drive is an extensible storage solution compared to SURF Drive, as you can access multiple storage environments (e.g., local file servers, SURF Drive, etc.) from one common

				Read Only: Anyone	workspace. With SURF Research Drive you can add users from outside Dutch institutions and can use more than 500GB.
Yoda	Negotiable €	✓	✓	Anyone at any Organization	Yoda (Your Data) is a research data management service that enables Utrecht University researchers to securely deposit, share, publish, and preserve research data during all stages of a research project.
O: Network Drive	Negotiable €	✗	✗	Only UU Geosciences	The O drive can only be used for sharing administrative data in the departments.
U: Network Drive	4 GB (default)	✗	✗	None	The U: This option will be phased out in the middle of 2025. So it is highly recommended

€: Usage may incur a cost, contact the Geo Data Team to discuss the costs

4.2 Research Data Lifecycle