



# Research Data Management

Neda Norouzi – Data Steward [n.norouzi1@tue.nl](mailto:n.norouzi1@tue.nl)

Linda Chen-Zijlstra – Research IT consultant [y.y.chen.zijlstra@tue.nl](mailto:y.y.chen.zijlstra@tue.nl)

# Who Are We?



# We are...



**Neda Norouzi**

Data Steward



**Linda Chen – Zijlstra**

Research IT consultant

# Today

## 09:00 – 12:00

09:00– 10:15 – RDM training

10:15 – 10:30 – Break

10:30 – 11:10 – Security and Research IT Resources

11:15 – 11:45 – How to package your data (a light version)

11:45 – 12:00 – Questions and discussion

# We hope after this workshop you will:

- ☐ Be able to evaluate your own research data management practices during the research data life cycle.
- ☐ Learn how to apply strategies to make your research data Findable, Accessible, Interoperable and Reusable (FAIR).
- ☐ Learn how to package your research data.
- ☐ Learn about security.
- ☐ The resources that you can use at TU/e (computational power, software, etc).
- ☐ Have fun!

# Ice breaker

**Talk to your neighbor and ask the following:**

- Why are you attending this workshop?
- What do you expect to learn today?



Photo on [iStock](#)

# Ice breaker

## Let me also ask you:

- Have you heard about RDM or any RDM elements, such as a DMP or Ethical approval?
- Are there any challenges you have faced in managing research data?



Photo on [iStock](#)

# What Is RDM?





# What is RDM?

## Research data management (RDM):

Concerns the way you collect, analyze, store, share, archive and publish research data, to satisfy the needs of current and future data users.



# Why RDM?



# Why Should I Care About RDM?

## External motivators:

TU/e wants you to!

- [TU/e RDM Policy](#) : [DOI](#)
- [TU/e Code of Scientific integrity and Conduct](#)
  - Ensuring scientific integrity, compliance, and responsible research

Laws and regulations

- [GDPR](#), Intellectual Property Law, WMO, and other legal obligations
  - how to safely store and transfer personal data
  - how to assess the privacy risks and mitigate those risks
  - Mishandling data can lead to Fines, Reputational Damage, Exclusion

## Funder Requirements

- No RDM = No funding – funders actively check compliance
- Publicly funded research should be transparent and accessible



**If you mess up, it affects us all!**

**Beware!**



**Scare Tactics**

# What does this Policy mean to me?

Requirement	Description	Resources
<b>Data Management Plans (DMPs)</b>	<p><b>For Research Staff (other than PhD and EngD)</b> A Data Management Plan is required for all research projects starting from September 2025.</p> <p><b>For PhD Candidates</b> A Data Management Plan is required for all research projects starting from September 2025. The DMP must be added to the <a href="#">Hora Finita</a> system at the 9-month evaluation and in the PhD defense package.</p> <p><b>For EngD (change to be approved by the UCC)</b> A Data Management Plan is required for all research projects starting from September 2025.</p> <p><b>For BSc and MSc</b> A Data Management Plan is only required for research projects that include data from humans.</p>	Research Cockpit
<b>Ethical Approval</b>	Ethical review is mandatory for all research involving human participants and/or processing personal data. See the Ethical Review Board (ERB) for more information.	Research Cockpit

[RDM Policy DOI](#)

# What does this Policy mean to me?

Requirement	Description	Resources
<b>Secure Data Storage</b>	All research data generated or processed at TU/e must be stored in university-provided solutions (central or departmental/research group storage).	Research Cockpit
<b>Data Agreements</b>	Mandatory data agreements for (personal or IP-protected) data sharing with external parties.	Research Cockpit
<b>FAIR Data</b>	Final research data must follow the FAIR (Findable, Accessible, Interoperable, and Reusable) principles and deposited in a data repository, where required.	FAIR data guidelines
<b>Data Package</b>	Research data and documentation from published peer-reviewed articles must be archived in a data package.	Research Cockpit

[RDM Policy DOI](#)

# Why Should I Care About RDM?

## Internal motivators:

- ✓ Reduce the risk of data loss
- ✓ Helps comply with TU/e procedures
- ✓ Makes research results more accessible
- ✓ Increases quality of data
- ✓ Increase efficiency – save time and avoid last-minute stress
- ✓ Get recognition – data citations and code re-use boost your impact
- ✓ Enable collaboration – through data sharing
- ✓ Increases grant funding opportunities
- ✓ Last but not least, a good practice for your career in science

💡 **A little effort upfront will save you a lot of trouble later!**

## Benefits for you



**It is not all bad!**

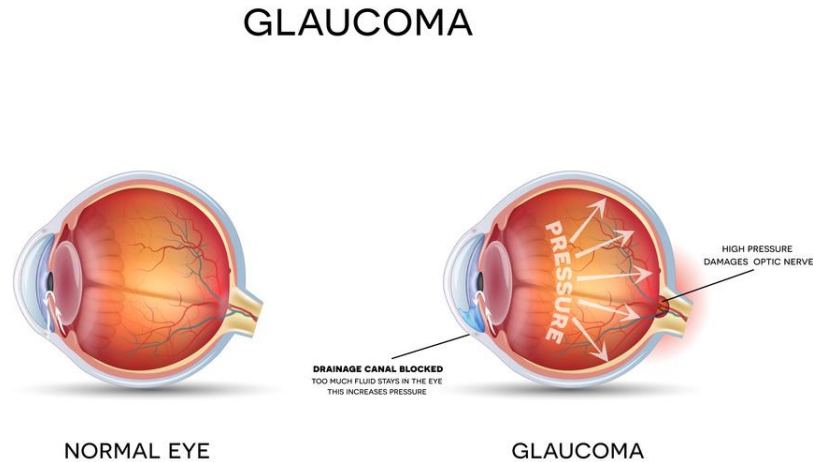
# RDM in Research Data Lifecycle(RDL)



# Example Research Project

**Research Question:** What causes variation in the rate of development of glaucoma in patients over the age of 50?

- Collaboration between: Industry, hospital and TU/e
- Dataset
- Code (R-script)





# Assignment 1

What are research data management topics you need to consider with glaucoma example?

- Glaucoma example
- Your own research project



# Assignment 1

What are research data management topics you need to consider with glaucoma example?



# Assignment 1

What are research data management topics you need to consider with glaucoma example?

- Ethics
- Privacy
- IPR
- Agreements
- Collection
- Sharing
- Storage
- DMP
- FAIR
- Documentation
- Folder structure
- Filenames
- Encryption
- Anonymization
- Versioning
- Metadata
- Software
- Archiving
- Publishing
- Repositories
- Licensing



# Data Management Plan(DMP)



# Plan ahead – good data management plan (DMP)

- Check TU/e RDM guidelines and funders requirements
- Consider how much data you will generate
- Select appropriate data storage
- Which tools you will need
- How you will organize and document data
- Consider costs for research data management
- When applicable, ask for ethical approval and assess privacy risks



<http://doi.org/10.5281/zenodo.3332807#>

# What is DMP? And why it is important

DMP is a living document that outlines how you will handle your research data during a project

A DMP helps you:

- ➔ at an early stage to plan how to make your research data **FAIR** and your research reproducible
- ➔ increase the impact of your research and make it more efficient

A typical Data Management Plan includes the following key elements:

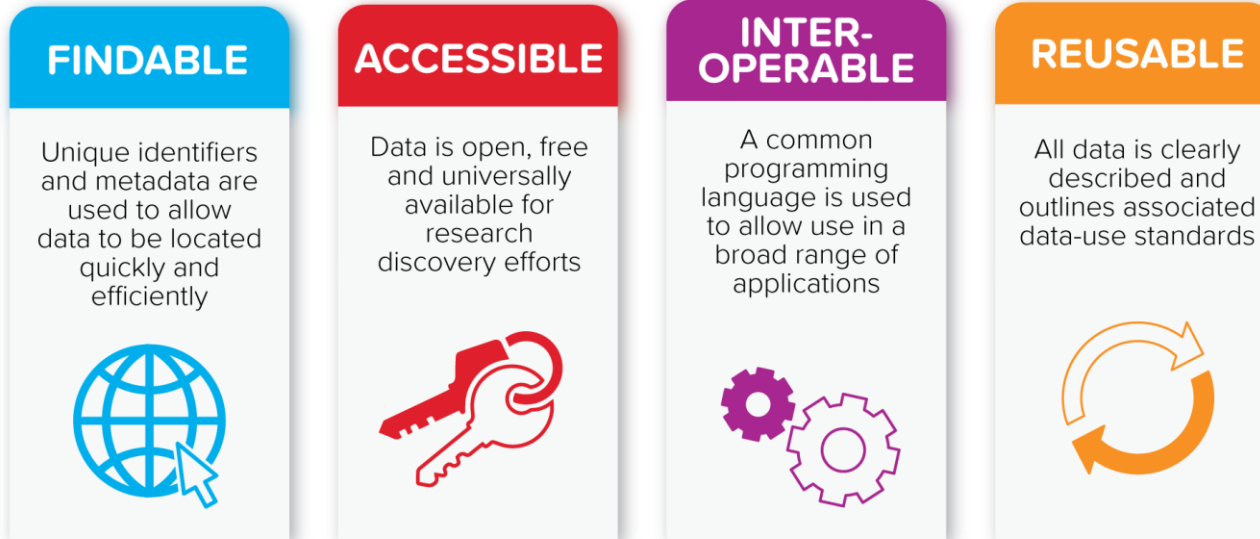
- Data Description
- Data Documentation and Metadata
- Data Storage and Backup
- Data Preservation and Archiving
- Data Sharing and Accessibility
- Data Security and Ethical Considerations
- Intellectual Property Rights
- Data Management Resources

# FAIR Principles



# FAIR

Research data should be managed or taken care of with the FAIR principles in mind – aimed at facilitating re-use of research data





# Making your data FAIR has several benefits:

- You increase the findability of your work and data, which leads to more citations.
- Your research can be reproduced easily, and it becomes more transparent, which further improves your research reliability.
- It helps other researchers (and your future self) understand the data and the research project.
- When your research is easily accessed, it enables more collaboration during your research and facilitates new collaborations in the future.
- When new collaborations emerge, it allows you to understand other perspectives and other research questions, increasing the impact of your data and work.

# FAIR

Research data should be managed or taken care of with the FAIR principles in mind – aimed at facilitating re-use of research data

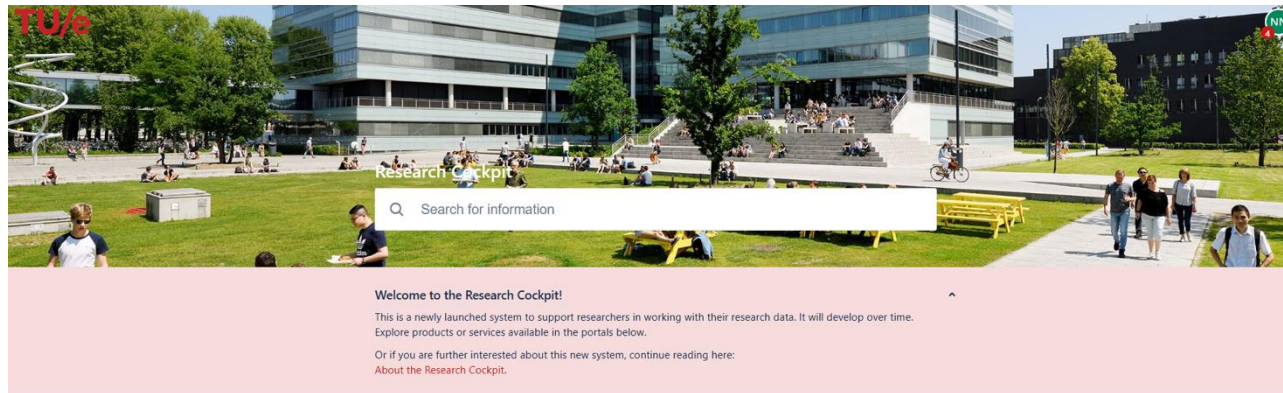


# How to make a DMP?



# Is there any Template?

- TU/e DMP(approved by various funders) → [Research Cockpit](#)



## Topics

### Plan your research

- Data Management Plan (DMP)
- Ethical Review
- Research Data Classification Guidelines

### Do your research

- I want to make use of GitHub
- Create a new anDREa workspace
- Cryptomotor Hub: introduction

### Share your research

- Repositories

## Featured portals



### GitLab

A comprehensive DevOps platform designed to facilitate the entire



### anDREa

A secure environment for processing (high-confidential) data.



### GitHub

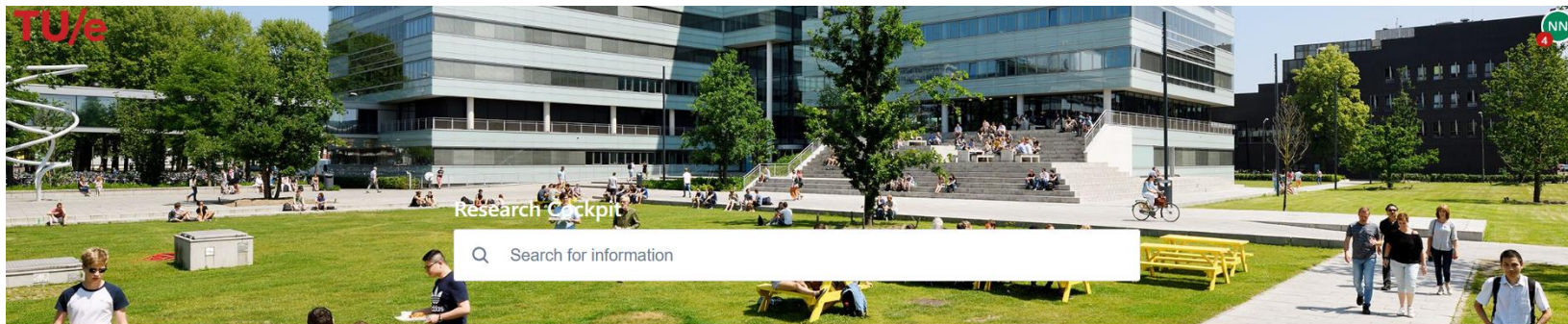
Collaboratively develop and share your software on the GitHub

# What you can find on Research Cockpit?

- The Research Cockpit is the system to create your Data Management Plan, also apply for Ethical Review.
- With the assistance of your departmental data steward, you will then be guided through the processes and data requirements for your research project.
- TU/e DMP has been approved by various funders.
- You can request various resources via Research Cockpit.
- A collaborative effort between various support teams with involvement of researchers.

Link to Research Cockpit <https://cockpit.research.tue.nl/helpcenter/research/>

# DMP process in Research Cockpit



## Welcome to the Research Cockpit!

This is a newly launched system to support researchers in working with their research data. It will develop over time. Explore products or services available in the portals below.




Or if you are further interested about this new system, continue reading here:  
[About the Research Cockpit.](#)

### Topics

#### Plan your research

-  Data Management Plan (DMP)
-  Ethical Review
-  Research Data Classification Guidelines

#### Do your research

-  I want to make use of GitHub
-  Create a new anDREa workspace
-  Cryptomotor Hub: introduction

#### Share your research

-  Repositories

### Featured portals



#### GitLab

A comprehensive DevOps platform designed to facilitate the entire



#### anDREa

A secure environment for processing (high-confidential) data.



#### GitHub

Collaboratively develop and share your software on the GitHub

# DMP process in Research Cockpit

Choose type of project

Research project




 Researchers

Create a Data Management Plan

- For scientific purposes:
  - my own research
  - funding
  - collaborations

Student project



 Students

Create a Data Management Plan

- For educational purposes:
  - part of a course
  - bachelor thesis
  - master thesis



## Data Management Plan (DMP)

A DMP is a document that describes how your research data will be handled throughout the project life cycle. If you have any questions or you are not sure if you need to do a DMP, the [data steward form your department](#) will be happy to help you.

### Guided tour of the DMP & ERB interface

Watch the below video to see how the DMP can be created and get introduced to key functionalities of this system:



Clicking "Create a Data Management Plan" will guide you through registering your project. This process will then automatically generate a DMP tailored to your needs.





# Research Cockpit

**Data Management Plan & Ethical Review Board Approval**

# TU/e DMP includes...

- **Basic Project information**

- Who is involved
- When will you collect data
- Funder information

- **Data**

- Source
- Size
- Study population
- Type of data – both personal and measurement data
- Anonymization

- **Data handling**

- Sharing
- Processing tools
- Storage
- Publishing

- **Privacy checklist**

# How to handle Personal Data?



**“Any information  
relating to an  
identified or  
identifiable natural  
person”**

**or in other words**

**Any data that  
can be directly  
or indirectly  
traced back to an  
individual**

## What is personal data?

### Identifying

Name  
Username  
Unique identifier  
Contact details



### Medical & health

Medical -  
conditions  
Medication use  
Mental health  
Genetic data



### Physical Characteristics

Height  
Weight  
Sex  
Age



### Biometric

Pictures  
Videos  
Fingerprints  
Audio recordings



### Professional

Education  
Employment



### Sexual information

Sex life  
Sexual -  
preferences  
Gender identity



### Demographics

Income brackets  
Age range  
Nationality



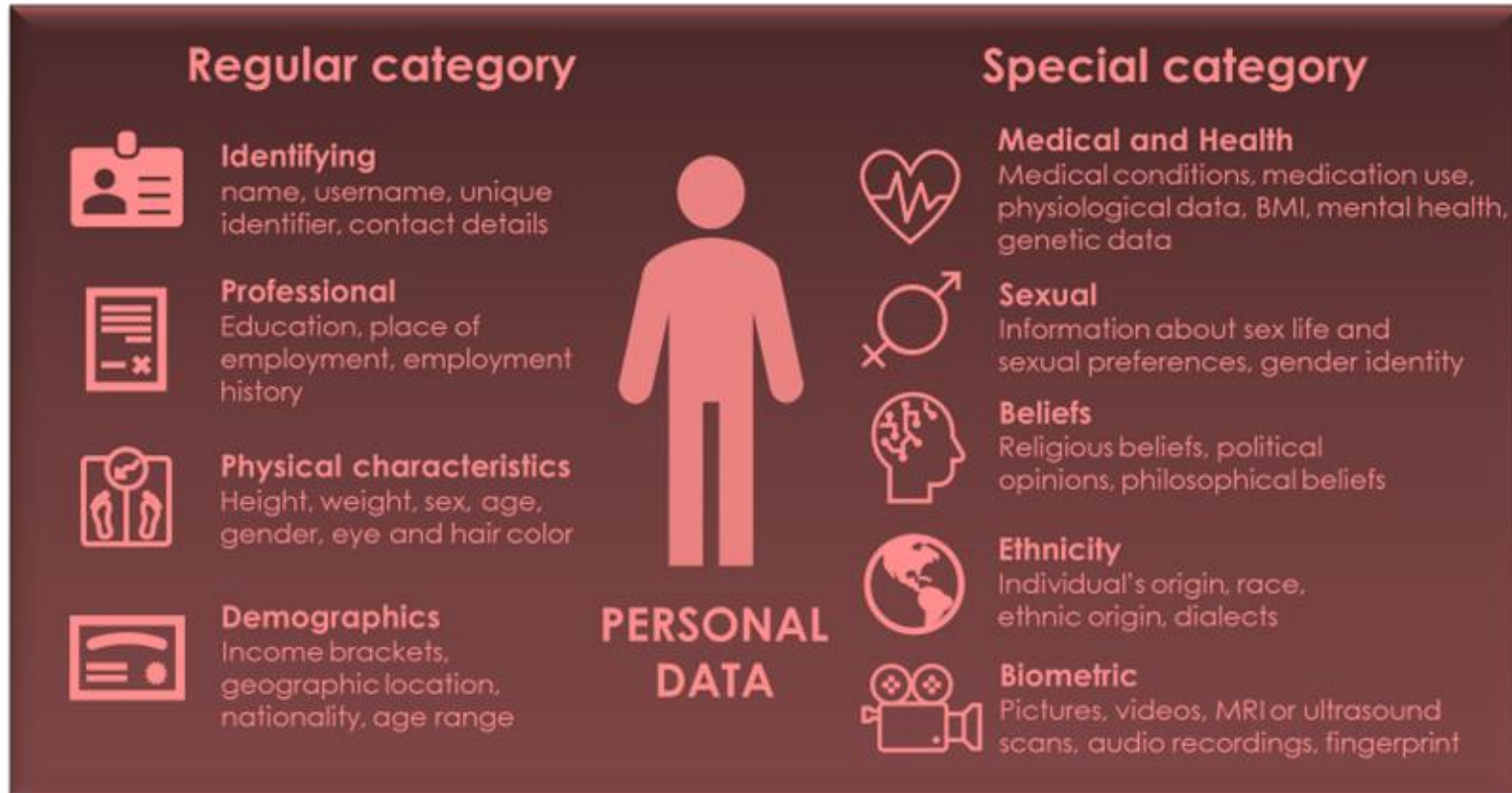
### Beliefs

Religious beliefs  
Political opinions  
Philosophical  
Beliefs

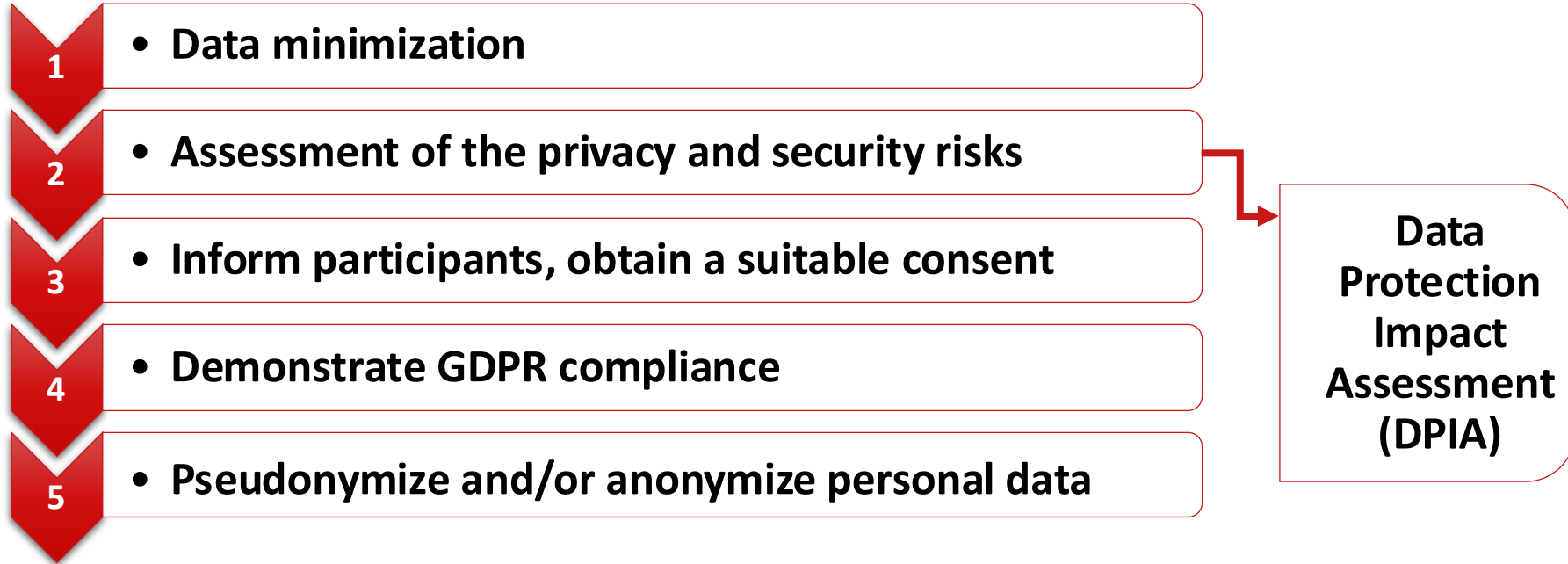




# Personal data



# Best practices: protecting privacy of participants



# Ethical & privacy review at TU/e

Mandatory for all research or educational projects where human participants, or personal data is involved.

Aim: to protect human participants and/or their personal data from harm

➤ You are responsible!

## *Risk levels and processing time:*

**Low** and **medium** risk studies are assessed by the departmental ERB. The processing time is **two working weeks**.

**High-risk** studies are assessed by the central ERB. The processing time is **six working weeks**.

# Ethical review process at TU/e

1. Choose type of project
2. Fill in a Project Registration form
3. Complete a Data Management Plan (DMP)
4. Approved DMP
5. Complete ERB Form

Lucia Forrová raised this on 2024-09-18 12:40 [Hide details](#)

Ethical Review Board [OPEN FOR EDITS](#) ... ^

### Ethics Review Template

Via this template you can create (or upload) your Ethics Review and request a review by the Ethical Review Board (ERB). This Ethical Review template should be completed for every research study that involves human participants or personal data.

**Questions:** If you have questions for the ethics committee, add them at the end of the form as a comment under "Activity".

**Submit:** Once you've completed the form, click "Submit". Your form will be automatically sent to the Ethics Review Board for revision.

1. Project Title \*

My Project

2. Are you a? \*

☐ Bachelor student

☒ Master student

☐ Researcher (PhD, academic staff)

☐ Other

3. Name of the researcher (Student or Researcher) \*

John Doe

4. Email of the researcher (Student or Researcher) \*

j.doe@tue.nl

Status

**DRAFT**

Notifications on

Retract my ERB request

Request type

Ethical Review Board Request

Shared with

Lucia Forrová  
Creator

Sandor Schmikli

+ Share

Issue Links

relates to (1) ^

DMP-304 My Project

**REVISED (POSITIVE ADVISE)**



# Advantages of DMP & ERB in Research Cockpit

- Simplified forms
- Status updates and progress notifications
- Automated background processes
- Automatization rules
- Fill in only what is relevant for you
- Nice overview progress of your application
- Submit and it will go to the right person
- Faster (automatic) approval

# Data Organization & Documentation



# Examples of documentation

- **README** files that provide information about associated data files, data formats, variables, etc.
- Your **DMP** describing how you intend to manage data throughout your project.
- **Documentation** about **equipment** used in your project.
- **Project plans or protocols** detailing your methods.
- **Code books** outlining procedures for coding and categorizing data.
- **e/Laboratory notebooks** describing your projects' progress, on paper or in electronic format.
- **Software syntax** that dictates how commands are structured and written in a programming language or software application.
- Documentation referring to **ethical or legal approvals** obtained at the start of the project.
- **Informed consent forms**, demonstrating permission to collect personal data from human participants in your project.
- Any **contracts** made in the project, if you work with a company for example.

# Make your data human- and machine readable

- Describe your variables and give them clear descriptive names
- Tidy your (tabular) data

# Data organization

## 1. Name of a project

### 1.1. Raw data and metadata

#### 1.1.1. Raw data

#### 1.1.2. Metadata

### 1.2. Processing and analysis files

#### 1.2.1. Importable data files (working copy of original data)

#### 1.2.2. Command files (files containing the code)

#### 1.2.3. Analysis files (cleaned and processed data files)

### 1.3. Documents (codebook, readme file, final paper)

### 1.4. Literature

Name	Patient number	Age	Sex	Glaucoma stage	Glaucoma stage	Glaucoma stage
				start	1 month	3 months
Trevon Cisneros	GS4687	720	Male	1	1	2
Leonard Robbins	GS4297	624	Male	4	6	8
Johan Khan	GS3960	696	Male	6	7	8
Ralph Huang	GS4501	852	Male	2	2	2
Caiden Noble	GS7823	780	Male	1	3	6
Ailyah Chan	GS2649	684	Female	1	2	2
Calvin Bullock	GS5643	624	Male	3	5	5
Amari Sanchez	GS7294	984	Female	5	5	6
Hazel Villanueva	GS2876	888	Female	7	7	8
Esteban Paul	GS7413	672	Male	2	4	5
Lawson Leonard	GS3592	756	Male	5	6	8
Mara Ponce	GS2748	864	Female	4	5	7

# Tidy glaucoma data

Messy data

Patient number	Age	Glaucoma stage start	Glaucoma stage 1 month	Glaucoma stage 3 months
GS4687	720	1	1	2
GS4297	624	4	6	8

Tidy data

Patient_number	Age (months)	Glaucoma _stage	Time_point
GS4687	720	1	0
GS4687	721	1	1
GS4687	723	2	3
GS4297	624	4	0
GS4297	625	6	1
GS4297	627	8	3

# Make your data human- and machine readable

- Describe your variables and give them clear descriptive names
- Tidy your (tabular) data
- Convert your data to open, non-proprietary formats
- Describe your dataset in a readme file/metadata file for discovery (title, creators, description, etc. of the dataset)



## Metadata file

Data about the data, Machine readable

Publish Meta Data only record in

[4TU.ResearchData](#)

Store xml copy in the data package

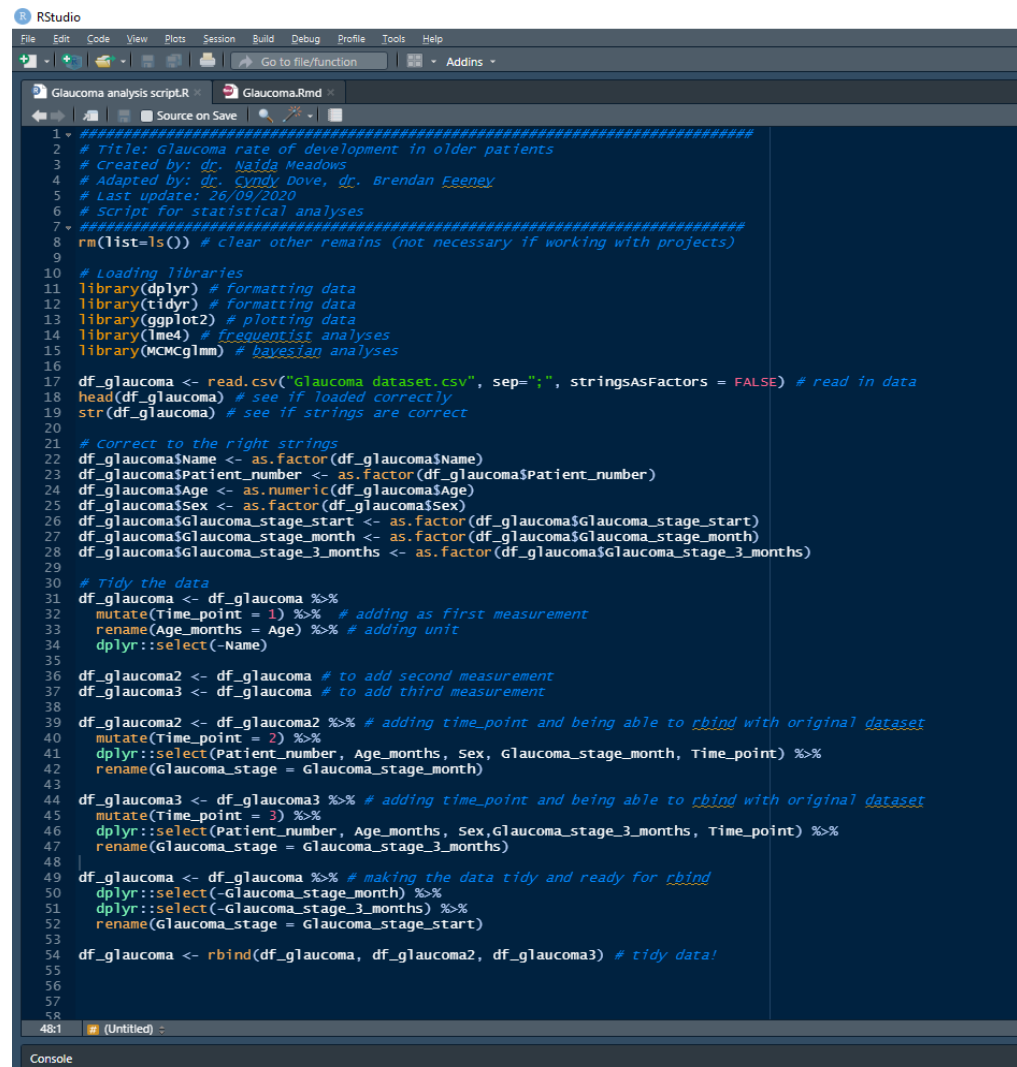
You can use:

You can use:

DCC Metadata Standards: [Disciplinary Metadata | DCC](#)

Research Data Alliance Metadata

Standards: <http://rd-alliance.github.io/metadata-directory/>



```
1 #####
2 # Title: Glaucoma rate of development in older patients
3 # Created by: dr. Naida Meadows
4 # Adapted by: dr. Cindy Dove, dr. Brendan Feenev
5 # Last update: 26/09/2020
6 # Script for statistical analyses
7 #####
8 rm(list=ls()) # clear other remains (not necessary if working with projects)
9
10 # Loading libraries
11 library(dplyr) # formatting data
12 library(tidyr) # formatting data
13 library(ggplot2) # plotting data
14 library(lme4) # frequentist analyses
15 library(MCMCglmm) # bayesian analyses
16
17 df_glaucoma <- read.csv("Glaucoma dataset.csv", sep=";", stringsAsFactors = FALSE) # read in data
18 head(df_glaucoma) # see if loaded correctly
19 str(df_glaucoma) # see if strings are correct
20
21 # Correct to the right strings
22 df_glaucoma$Name <- as.factor(df_glaucoma$Name)
23 df_glaucoma$Patient_number <- as.factor(df_glaucoma$Patient_number)
24 df_glaucoma$Age <- as.numeric(df_glaucoma$Age)
25 df_glaucoma$Sex <- as.factor(df_glaucoma$Sex)
26 df_glaucoma$Glaucoma_stage_start <- as.factor(df_glaucoma$Glaucoma_stage_start)
27 df_glaucoma$Glaucoma_stage_month <- as.factor(df_glaucoma$Glaucoma_stage_month)
28 df_glaucoma$Glaucoma_stage_3_months <- as.factor(df_glaucoma$Glaucoma_stage_3_months)
29
30 # Tidy the data
31 df_glaucoma <- df_glaucoma %>%
32   mutate(time_point = 1) %>% # adding as first measurement
33   rename(Age_months = Age) %>% # adding unit
34   dplyr::select(-Name)
35
36 df_glaucoma2 <- df_glaucoma # to add second measurement
37 df_glaucoma3 <- df_glaucoma # to add third measurement
38
39 df_glaucoma2 <- df_glaucoma2 %>% # adding time_point and being able to rbind with original dataset
40   mutate(time_point = 2) %>%
41   dplyr::select(Patient_number, Age_months, Sex, Glaucoma_stage_month, time_point) %>%
42   rename(Glaucoma_stage = Glaucoma_stage_month)
43
44 df_glaucoma3 <- df_glaucoma3 %>% # adding time_point and being able to rbind with original dataset
45   mutate(time_point = 3) %>%
46   dplyr::select(Patient_number, Age_months, Sex, Glaucoma_stage_3_months, time_point) %>%
47   rename(Glaucoma_stage = Glaucoma_stage_3_months)
48
49 df_glaucoma <- df_glaucoma %>% # making the data tidy and ready for rbind
50   dplyr::select(-Glaucoma_stage_month) %>%
51   dplyr::select(-Glaucoma_stage_3_months) %>%
52   rename(Glaucoma_stage = Glaucoma_stage_start)
53
54 df_glaucoma <- rbind(df_glaucoma, df_glaucoma2, df_glaucoma3) # tidy data!
55
56
57
58
59 48:1 (Untitled)
60 Console
```

# README file

“is a text file that introduces and explains a project. It contains information that is commonly required to understand what the project is about.”

“Because no one can read your mind”  
(<https://www.makeareadme.com/>)

Human readable

```
Documentation.txt - Notepad
File Edit Format View Help
NAME: Glaucoma rate of development in older patients
TYPE: Complete records for all patients developing glaucoma
SIZE: 400 patients, 10 variables (20Kb)

DESCRIPTIVE ABSTRACT:
For each patient in hospital A that has developed glaucoma, this dataset records the patient number, sex (male/female), age (in months), glaucoma stage and at which point the glaucoma was measured.

SOURCE:
Hospital A

METHOD:
Data was collected at Hospital A at the intake of the patient. The patient was informed about the collection of the data and the processing of the data for this research. Glaucoma stage was determined through measuring the pressure and inspection of the optic nerve. Data was subsequently stored in the patient's dossier, awaiting further statistical processing in R 3.3.1. Patients were asked about their lifestyle and various other vital metrics were record such as heartbeat, stress level, food intake, weight and height. The results of this research will be published and the pseudonymised data will be made available in a repository and deleted from all other devices at the end of the project.

VARIABLE DESCRIPTION:
Column
1 Patient Number: Unique number for each patient
2 Age: the age of the patient in months
3 Sex: the sex of the patient
4 Glaucoma stage: the severity of glaucoma (1 = none, 2 = barely, 3 = mild, 4 = moderate, 5 = half blind, 6 = badly, 7 = severe, 8 = thoroughly, 9 = extreme , 10 = blind)
5 Time_point: The time point at which the glaucoma in the patient was measured (0 = when coming into hospital, 1 = after exactly 1 month, 3 = after exactly two months)

Values are aligned and comma seperated. There are no missing values.

SPECIAL NOTE:
Patient GS2876 was measured at 1 month, 2 months and 4 months after being in the hospital.

List of abbreviations:
SD = standard deviation

Ln 3, Col 23 100% Windows (CRLF) UTF-8
```

# Make your data human- and machine readable

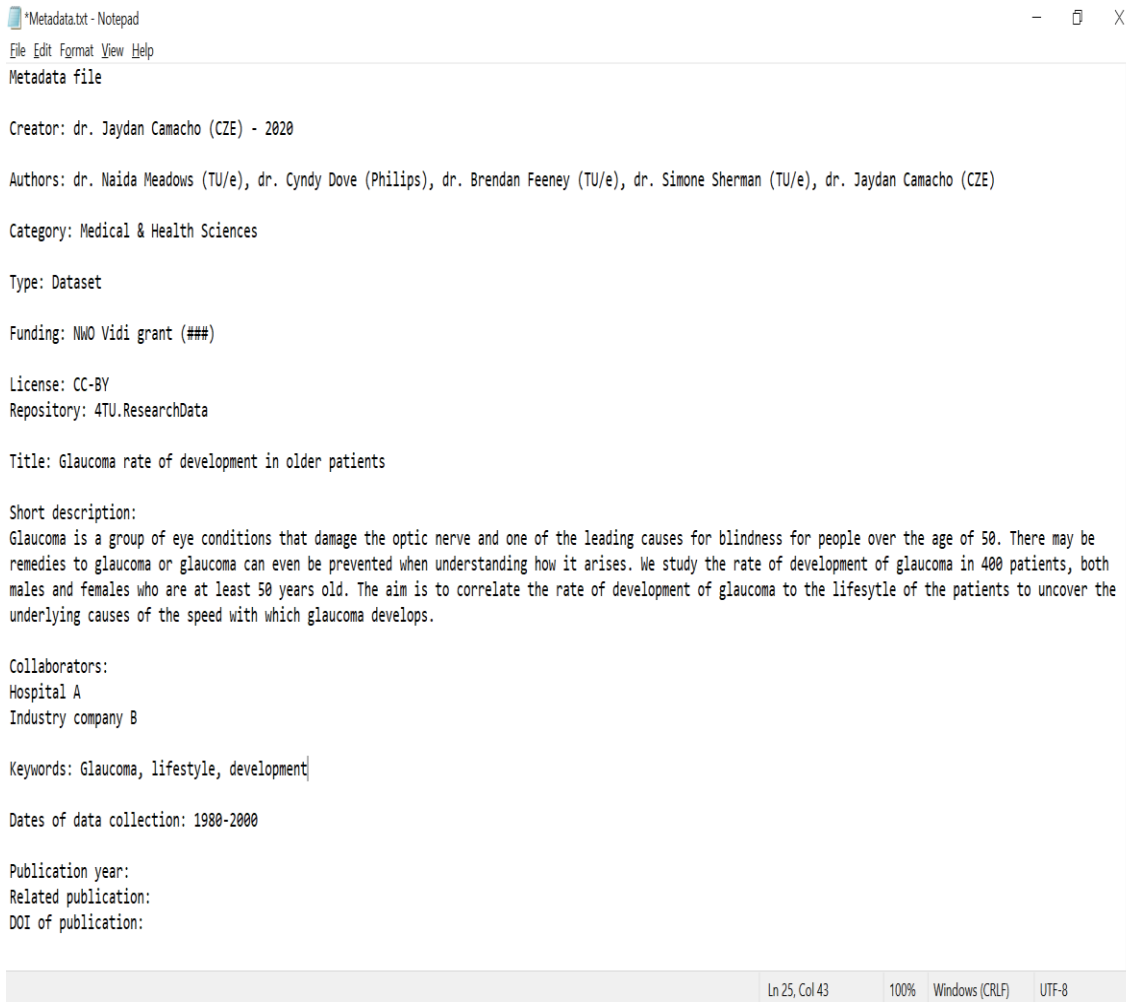
- Describe your variables and give them clear descriptive names
- Tidy your (tabular) data
- Convert your data to open, non-proprietary formats
- Describe your dataset in a readme file/metadata file for discovery (title, creators, description, etc. of the dataset)
- Write a documentation file/annotate code
- For code, once public add a license to your code (e.g. MIT/GPL) – must or copyright violation

# Documentation file

Stuff we need to know, but  
is not in any of the other  
files

File naming that makes  
sense

E.g. Project documentation,  
dataset documentation, the  
object description, code  
books, etc



The screenshot shows a Notepad window titled '\*Metadata.txt - Notepad'. The menu bar includes 'File', 'Edit', 'Format', 'View', and 'Help'. The text content of the file is as follows:

```
Metadata file

Creator: dr. Jaydan Camacho (CZE) - 2020

Authors: dr. Naida Meadows (TU/e), dr. Cyndy Dove (Philips), dr. Brendan Feeney (TU/e), dr. Simone Sherman (TU/e), dr. Jaydan Camacho (CZE)

Category: Medical & Health Sciences

Type: Dataset

Funding: NWO Vidi grant (###)

License: CC-BY
Repository: 4TU.ResearchData

Title: Glaucoma rate of development in older patients

Short description:
Glaucoma is a group of eye conditions that damage the optic nerve and one of the leading causes for blindness for people over the age of 50. There may be remedies to glaucoma or glaucoma can even be prevented when understanding how it arises. We study the rate of development of glaucoma in 400 patients, both males and females who are at least 50 years old. The aim is to correlate the rate of development of glaucoma to the lifestyle of the patients to uncover the underlying causes of the speed with which glaucoma develops.

Collaborators:
Hospital A
Industry company B

Keywords: Glaucoma, lifestyle, development

Dates of data collection: 1980-2000

Publication year:
Related publication:
DOI of publication:
```

The status bar at the bottom indicates 'Ln 25, Col 43', '100%', 'Windows (CRLF)', and 'UTF-8'.

# Data Processing



# During research

## Analysing your data

- Versioning your data and documentation

File naming conventions: Systematic file names will help you identify the correct datasets without having to open them.

- How to retrieve files
- Relevant components
- Reasonable length
- No special characters or space
- Unite workflow in your research group



# Where to Store Research Data?



# TU/e-recommended Storage Options

Option	Security	Scalability	Self-service access management
Network Drive	🏆 Best: Usable for special category personal data	🏆 Best: Max. 100 TB	❌ Top folder: No (only via servicedesk) ✅ Subfolders: Yes
SURF Research Drive	✅ Good: EU based	✅ OK: Max. 10 TB	✅ Yes
MS Teams / SharePoint	✅ Reasonable: US based	👉 OK: Max. 2.5 TB	✅ Yes
MS OneDrive	✅ Reasonable: US based	😐 So/so: Max. 1 TB	✅ Yes (but only one owner!)



# Where to Store my Research Data?

- Ensure the system makes back-ups!
- Do not put data on personal devices
- Hard-drive encryption (e.g., Bitlocker)

- For more information about the available storage solutions at TU/e, visit [PAR Solution Searcher](#).

- TU/e Research Data Management: [Website](#) & [Intranet](#).

<u>Data Classification (Confidentiality)</u>	Examples of Data	Student	Staff - only within TU/e	Staff - external collaboration
Low	Non-personal data	One Drive	Network Drive/Teams	SURF Research Drive/Teams
Medium	Personal data	OneDrive	Network Drive/Teams	SURF Research Drive/Teams
High	Special categories of personal data	One Drive + <a href="#">encryption</a>	Network Drive	SURF Research Drive + <a href="#">encryption</a>

This guide applies only to research data. For non-research data (e.g., administrative or education data), contact the respective [data domain coordinators](#).

# How to Share Research Data?



# Sharing data

## Sharing data

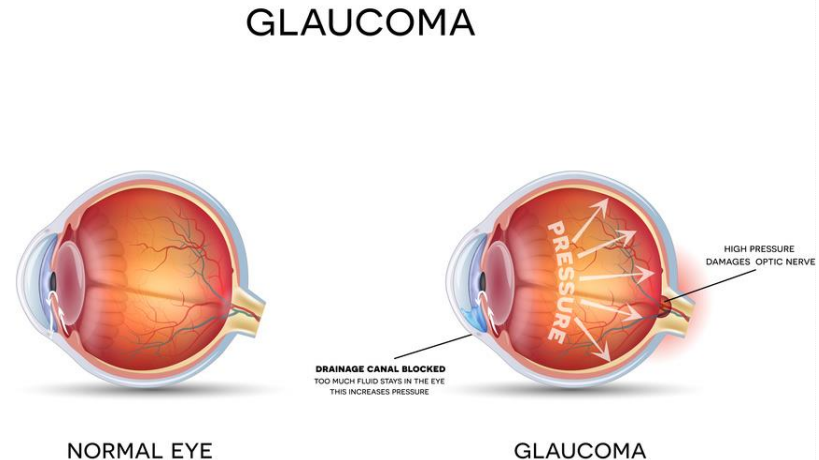
- Surffilesender: any data (only transfer)
- Network drive (sharing internally): research data (incl. storage)
- ResearchDrive (sharing externally): research data (incl. storage)
- Microsoft Teams /Sharepoint: office data

If used appropriately (e.g., pseudonymized) → personal data

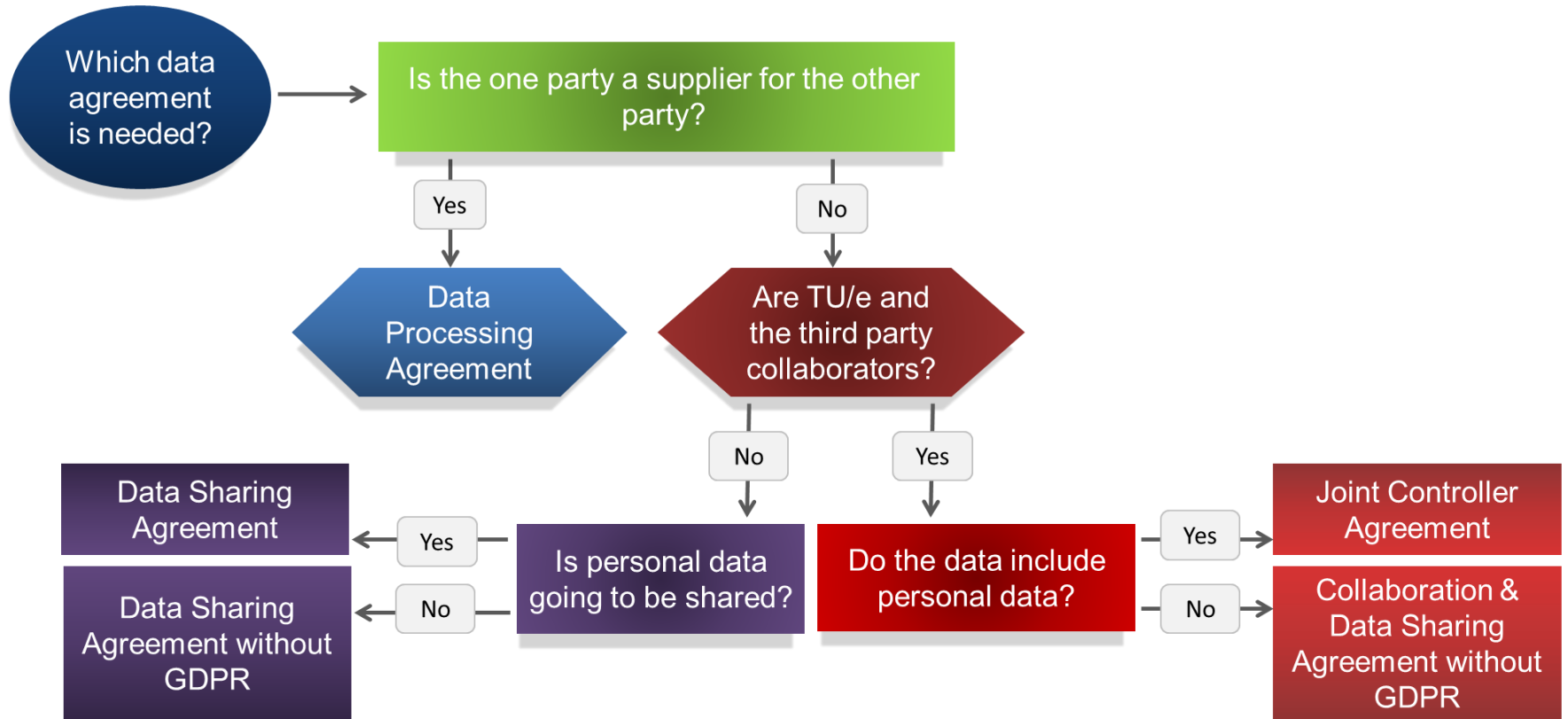
# Storage Glaucoma example

Where would you store the glaucoma data at TU/e?

- A) SurfDrive
- B) Onedrive
- C) ResearchDrive
- D) Network drive



# Sharing data with collaborators



# How to Publish Research Data?



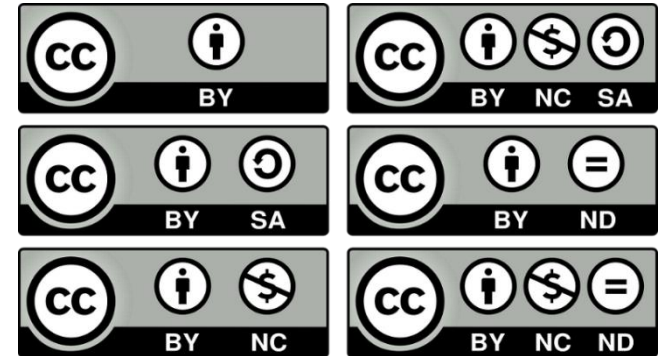
# Data repositories

- Depositing data/code is a must for long-term storage
- Choose a repository where other researchers of your discipline are sharing their data
- General repository (4TU.ResearchData, Zenodo)
  - Assigns DOI
  - Adequate metadata required
  - Select user license
- Connect your dataset with your publication (and vice versa)



# Licences

- Data licensing
  - [Creative Commons \(CC-BY-NC-SA-ND-0\)](#)
  - Open database licenses
  - Once depositing in a repository
- Software licensing
  - MIT
  - GPL
  - Once public (on GitHub/Gitlab)





# Publishing data

**“as open as possible, as closed as necessary”**

Valid reasons not to make your data open

1. Confidentiality / privacy
2. National security
3. Intellectual property



# Where to Archive Research Data?



# RAPS\_Reserach Archival Package Solution

RAPS is a:

- Closed TU/e archive for data underlying an accepted journal article
- Ensures long-term preservation & research integrity

**Before requesting RAPS:**

- Register article in Pure
- Prepare data package following [RAPS guidelines](#)

# Archiving vs. Publishing



# Archiving vs Publishing

## When to Use RAPS (Archiving)

- Article is accepted in a peer-reviewed journal
- Store the underlying research data securely
- Ensure long-term preservation
- Data is not necessarily publicly accessible
- Used for institutional compliance and record-keeping

## When to Use a Repository (Publishing)

- Data needs to be openly accessible to others
- You want a DOI for citation and reuse
- Required by funders or journals (FAIR requirements)
- Goal is sharing and enabling reuse of data

# How can Data Stewards help?



# How can data stewards help?



- Advise or help on RDM-related questions: [n.norouzi1@tue.nl](mailto:n.norouzi1@tue.nl) or [rdmsupport@tue.nl](mailto:rdmsupport@tue.nl)
- Privacy & Ethics
- Data Management Plans
- Metadata and documentation
- Data sharing/storage
- Archiving
- Making data FAIR

## AND MUCH MORE!

More information on RDM:  
<https://www.tue.nl/en/our-university/library/rdm/>

# More information on RDM:

- [TU/e RDM Website](#)
- [RDM Handbook](#)
- [RDM Booklet](#)
- [TU/e Research Data Solutions Guidelines](#)





DATE

29/01

## Introduction to Research Data Management

<https://doi.org/10.5281/zenodo.18479954>



DATE

26/02

## Research Archival Package Solution (RAPS) - preservation of data underlying scientific

<https://doi.org/10.5281/zenodo.18801456>



DATE

26/03

## Data Management Plan and why every project needs one

<https://doi.org/10.5281/zenodo.19694418>



DATE

30/04

## Privacy, Security and Ethics in Research Projects

<https://doi.org/10.5281/zenodo.20050309>



**DATE**  
28/05

## How to create a data package

[Register here](#)



**DATE**  
25/06

## Version Control in Action

[Register here](#)



**TIME**  
13:00 - 13:30



**LOCATION**  
ONLINE - TEAMS

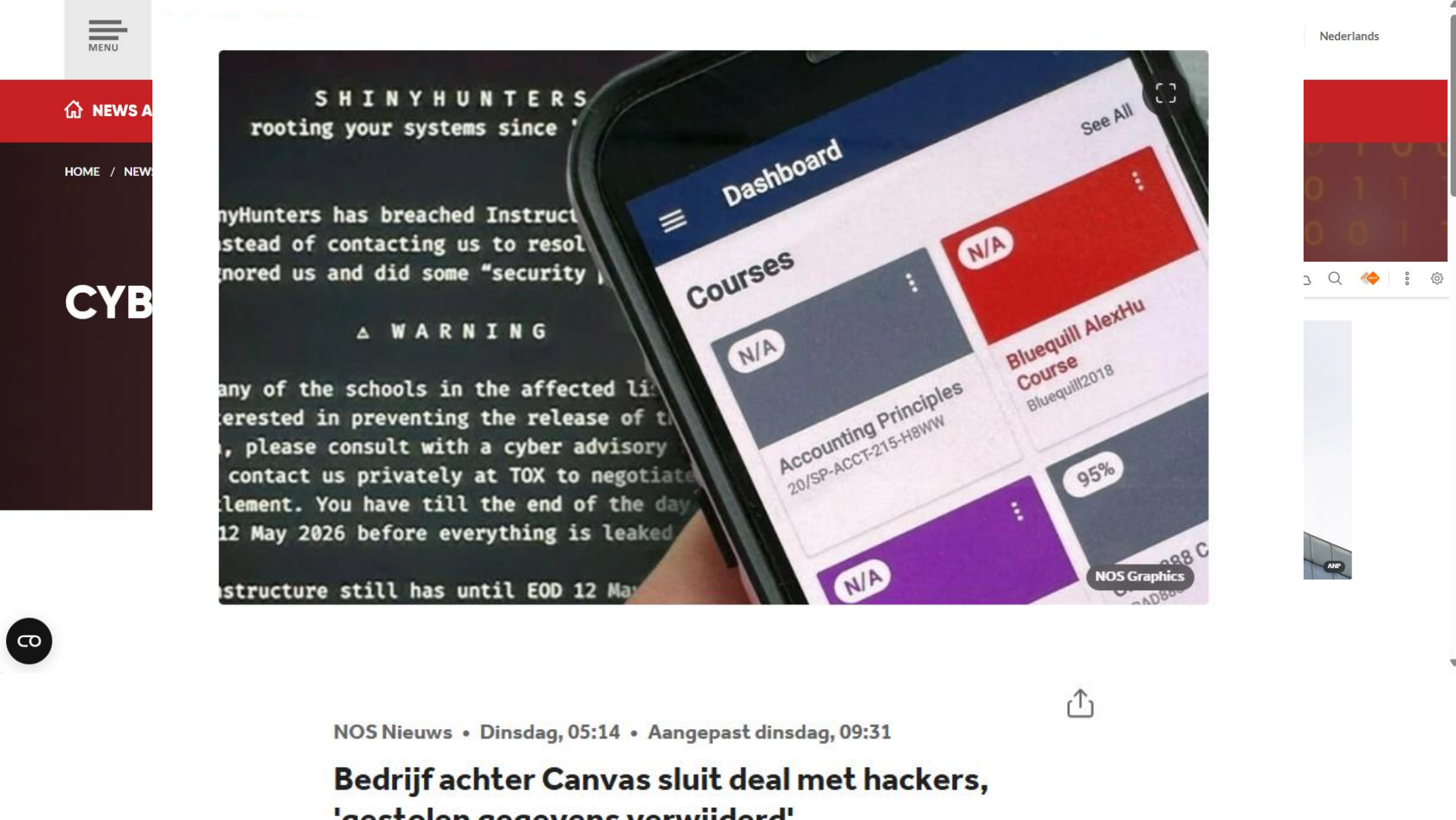


RDMSUPPORT@TUE.NL

# Break

# Security (a brief introduction)





SHINYHUNTERS  
rooting your systems since

nyHunters has breached Instruct  
instead of contacting us to resol  
ignored us and did some "security

△ WARNING

any of the schools in the affected li  
interested in preventing the release of t  
, please consult with a cyber advisory  
contact us privately at TOX to negotiate  
lement. You have till the end of the day  
12 May 2026 before everything is leaked

structure still has until EOD 12 May

Dashboard

Courses

N/A

Accounting Principles  
20/SP-ACCT-215-H8WW

N/A

Bluequill AlexHu  
Course  
Bluequill2018

95%

N/A

NOS Graphics

NOS Nieuws • Dinsdag, 05:14 • Aangepast dinsdag, 09:31

Bedrijf achter Canvas sluit deal met hackers,  
'gestolen gegevens verwijderd'



# Why is information security important?

Information security is important for several reasons:

- Protection of Data
- Privacy Preservation
- Business Continuity
- Intellectual Property Protection
- Fraud Prevention
- Trust and Reputation
- Compliance Requirements
- Evolving Threat Landscape



# How is information security managed at the TU/e?

- Security Operations / Chief Information Security Officer
- But we also need your help!
  - CERT
  - Security awareness program
- More information?
  - Visit the privacy and security intranet page!
  - Policies and guidelines.

Contact CERT:  
Call (040-247) 5678 or  
Email [cert@tue.nl](mailto:cert@tue.nl)

- **Come to the Hub vector 2.203**

# Start from something simple...

## Research Data storage

- ✓ TU/e network storage and Research Drive for data,
- ✓ One Drive for documents/ presentations.
- ✓ myDRea/ SANE for sensitive data.

**X External HD and USB sticks are not recommended for research data.** Because it can be corrupted, like any other piece of hardware. They can be susceptible to fire and water damage, can overheat, and can malfunction. They can also be stolen....



## Personal and research information (Cyber-crime, data breach...)

- Report a possible data breach (Topdesk- IT services - Security – Report a risk).
  - Lock your computer when you leave your desk.
  - Use a security cable for laptops, available via the Self-Service Portal.
  - Shut down your PC entirely when it is not in use (end of workday, weekend, holiday)
  - Make use of a password safe, for your passwords.
  - Use **SURF file sender/ Cryptomator Hub** for sharing sensitive data to bigger groups.
  - Update your device shortly after you received a notification.
  - Make a backup of your documents or use a verified platform that automatically makes backups (TU/e network drive makes backups for you already).
  - Create a strong password, and keep your password to yourself.
- 81 Always choose for Multi factor Authentication (MFA).

# WHAT'S WRONG IN THIS PICTURE?



# AI tools guidelines for research

- Do not put in any privacy sensitive and/or confidential data  
Example: “Help me summarize the student and course information of John Doe, whose student ID is 123456.”
- Keep your login credentials secured (always)!
- Do not use LLMs for any prohibited/ illegal activities
- Practice Data Minimization while using LLMs, for example, only use the information that you need or turn off the usage of chat history
- Inform your manager about LLM usage for work
- Using local LLMs is preferable

## For more information

### **Privacy E-learning Course: Privacy – The Basics**

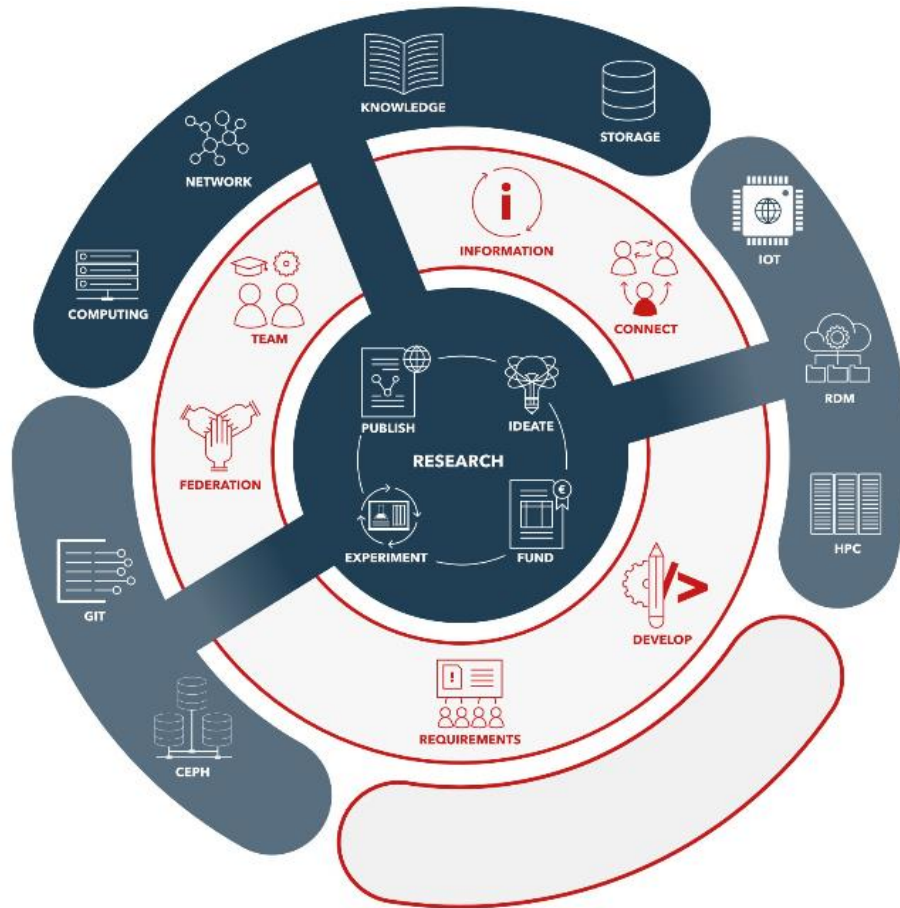
<https://tuenl.sharepoint.com/sites/intranet-privacy-security/SitePages/New-Privacy-E-learning-Course--Privacy-%E2%80%93-The-Basics.aspx>

[Privacy-security](#)

# Research IT Support



# Research IT consultant



- With a research background.
- We are sensitive to the latest IT solutions, the TU solution and research.
- Connect BME to LIS.
- Custom-made IT for research to help you answer your research questions.
- The aim is to make research more efficient, the IT infrastructure more sustainable for all researchers.

# How to find your information at BME and Hub Vector

**BME SharePoint** <https://tuenl.sharepoint.com/sites/intranet-biomedical-engineering>

**Hub Vector SharePoint** <https://tuenl.sharepoint.com/sites/intranet-biomedical-engineering/SitePages/HUB-Gemini.aspx>

# Resources we have in house...

- **Topdesk** [Home Page - Selfserviceportal TU/e \(topdesk.net\)](https://topdesk.net)
- **Research cockpit** <https://cockpit.research.tue.nl/helpcenter/research/>
- **Software** [TU/e Software - Selfserviceportal TU/e \(topdesk.net\)](https://topdesk.net)

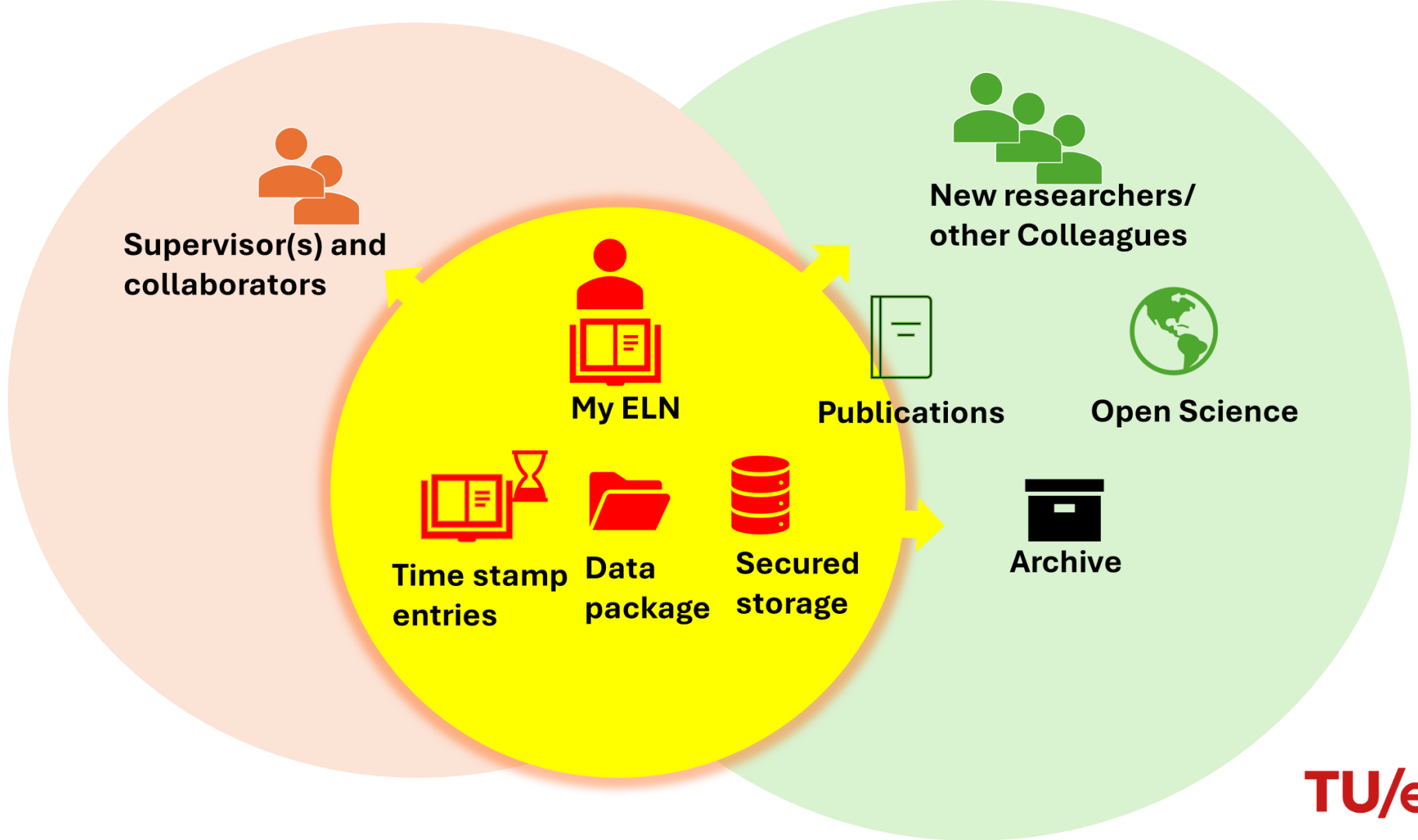




- **Research data storage** <http://storagefinder.tue.nl/>
  - Centralized storage (network drive)
  - Research drive
  - One drive
- **High-Performance computing**
  - HPC Umbrella cluster, HPC ondemand, **mate.cpu.q** and **bme-gpu2.q**
  - Surf Snellius, or research Cloud
  - **myDREa/ SANE** (sensitive data upload, short term storage, processing and sharing)
- **AI tools**
  - Copilot
  - ASReview
- **Hub Gemini BME Teams Channel – all welcome!**
  - For guidance, training announcements, extra documents, etc.

# Example of good practices in the lab

- elab journals (example 6 minutes)
  - Courtesy of Rachel Armstrong, Research group MBx, APSE, TU/e
- BiS ElectronicLabJournal example
  - Courtesy of Jan de Boer, BME, TU/e



Submission to data repository

RASP, 4TU research data

Traceable experiments  
and context

Data package

Storage and Archiving  
TU/e Network Drive

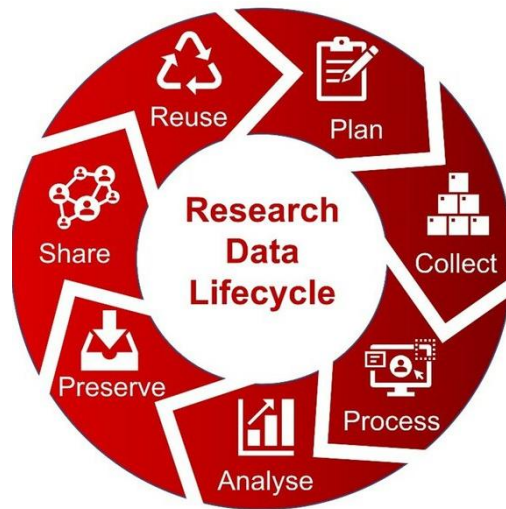
DMP

Research cockpit

Metadata standard

Data formats

Workflow, pipeline and  
computer resources



HPC/ Snellius

TU/e

# Good practices...

- Research data management – organised your data, store it properly, keep it clean and safe.
- elab journal
- Research software management <https://esciencecenter-digital-skills.github.io/research-software-support/>
- [The Turing way](#)

Find LIS on the TU/e Intranet pages: [LIS](#)

Hub Vector : Vector 2.203

BME intranet page: [BME](#)

## LIS Central

Contact: [LISservices@tue.nl](mailto:LISservices@tue.nl)

Telephone: 040-2472000 or 040-2475005 (Hub Vector)



[This Photo](#) by Unknown author is licensed under [CC BY](#).



# Library

- provides both library and archive services around the storage, preserving and availability to both academic and non-scientific information
- **Library Services**: provides a broad range of scientific content for the university and offers an inspiring physical library in the MetaForum building
  - Contact: [library.helpdesk@tue.nl](mailto:library.helpdesk@tue.nl)
- **Privacy**: supports TU/e employees with GDPR related questions and issues to make sure that personal (sensitive) data is handled safely
  - Contact: [privacy@tue.nl](mailto:privacy@tue.nl)

Bart van Overbeeke Photography



# A modern Literature review guide



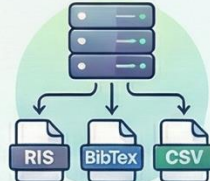
## Define Research Parameters

Outline clear research questions and identify relevant keywords, synonyms, and Boolean operators.



## Use Trusted Academic Databases

Start with TU/e Library Databases to ensure quality control and peer-reviewed results.



## Export Systematic Results

Save library search results in RIS, BibTeX, or CSV formats for tool integration.



## Prioritize with Active Learning

Use AI tools like ASReview to rank articles by relevance and reduce manual effort.



## Maintain Human Critical Reading

AI assists with sorting, but researchers must independently analyze methodologies and study limitations.



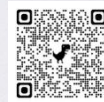
## Document the Workflow

Record search strategies and specific AI tools used to ensure full research transparency.

## TU/e Library Database & Search



## Using GenAI



## Visit the TU/e information Literacy Hub





# Open Access Support Team



[openaccess@tue.nl](mailto:openaccess@tue.nl)



ext. 2529



[www.tue.nl/openaccess](http://www.tue.nl/openaccess)

We can help you:

- ✓ Estimating Open Access publication fees in the budget proposal, during the grant application phase;
- ✓ Planning Open Science in your grant proposal;
- ✓ Finding optimal publication strategy compliant with Funder Open Access requirements;
- ✓ Finding the Green Open Access policy of a journal;
- ✓ Finding opportunities to publish Open Access free of charge or at reduced price;
- ✓ Uploading your manuscript in the TU/e repository or determining which another Open Access repository might be suitable;
- ✓ Organizing a workshop or a training in Open Access publishing;
- ✓ Any other questions related to Open Science

# Research data storage



## Storage – where to save my research data

	Data Sharing		Expandability	Remarks
	Within TU/e	external		
Network Drive	✓	X	✓✓✓✓	On-premise storage for large data set.
Research Drive	✓	✓	✓✓	Project base, the ownership can be transferred.
SharePoint	✓	✓	✓✓	Cloud platform for document management and content sharing for teams.
One Drive	✓	✓	X	Discontinue after the TU/e contract expires, export data to the network drive

# How to package your data (the light version)



# Data package at the end of your study, because:

- The supervisor or the team can continue with your work.
- Knowledge can be accumulated in the lab.
- It may end up with more publications or collaborations for your career.

**But the data need to be findable, and other people can understand the data**



To do this, we need to start organizing your data early on.

# Publishing vs. Archiving

Publishing: Publishing (including storage) the data is in a repository, by default open.

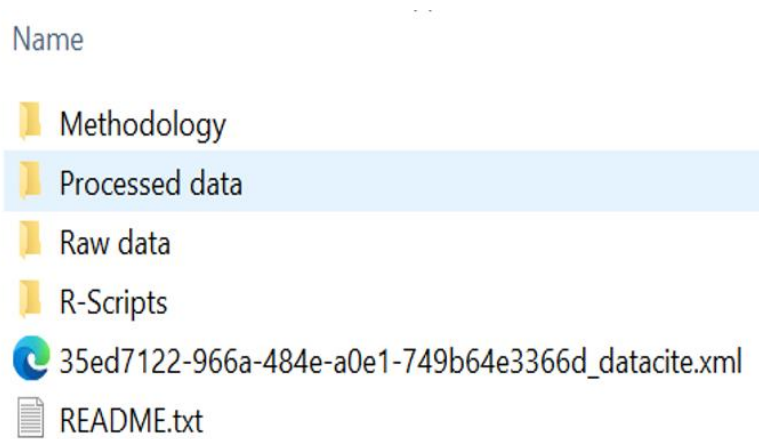
Archiving: Storing the data on a TU/e system (data curation is done by the TU/e), by default not open.

# Data package

The package is self-descriptive and contains the following information (metadata), read-only:

- Readme files OR Documentation (see template <https://doi.org/10.5281/zenodo.19692009>)
  - Create one readme file for each dataset
  - Name the file; not just readme, read\_me, ABOUT, etc.
  - Write your readme document as a plain text file
- Raw data files
- Processed data files (machine and human readable)
- Syntax files (Script/ code)
- Protocols (lab journals, references, SOPs)

# Data package example



- Clear file name and folders
- Metadata (4TU. Research Data)
- Local copy of GitLab
- Publications, non-data files
- Zip and store final version on a specific address
- Check! And make a personal copy (if allowed).
- Remove other copies from the TU/e systems.
- Use a similar template within a group.



# Data package example



# What to save and what not to save...

- ✓ Legal or legislative reasons (contract, funder policies..).
  - ✓ Meaningful results with metadata and well-documented (written in lab journals).
  - ✓ Do not make unnecessary copies (the TU/e Network drive made copies and backed them up for you already, with a disaster recovery function).
  - ✓ Be critical to your data, for example, microscopy data (crop the empty space), faster processing time.
  - ✓ Good habits can save you time and be efficient!
- 
- ✗ Backup of a backup is not needed if you store your data at the TU/e network Drive or Research Drive.
  - ✗ Test data, failed experiments.
  - ✗ Some intermediate data

**The full document for data package**  
**[Data\\_package\\_v2.3.docx](#)**

**Or join the data stewards workshops**

# Open Q&A

