Surrounded by Science

Learning Paths towards Science Proficiency

Research and Innovation Action in the European Union's Horizon 2020 Programme

Grant Agreement no. 101006349

Deliverable 1.2

Data management plan

Editor	Jakob Sikken (UT)
Date	31 March 2022
Dissemination Level	Public
Status	Final

The Surrounded by Science project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement no. 101006349. This publication only reflects the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

The Surrounded by Science Consortium

Participant No. *	Participant organisation name	Short name	Country
1 (Coordinator)	Universiteit Twente	UT	Netherlands
2	Ellinogermaniki Agogi Scholi Panagea Savva AE	EA	Greece
3	European Physical Society Association	EPS	France
4	Nuclio Nucleo Interactivo de Astronomia Associacao	NUCLIO	Portugal
5	Fondazione IDIS-Citta della Scienza	IDIS	Italy
6	The Lisbon Council for Economic Competitiveness and Social Renewal asbl	LC	Belgium
7	Weizmann Institute of Science	WIS	Israel
8	Norges Teknisk-Naturvitenskapelige Universitet	NTNU	Norway

Contributors

Name	Institution
Jakob Sikken	UT
Tessa Eysink	UT
Natasha Dmoshinskaia	UT

Legal Notices

Executive Summary

This deliverable describes the data management plan for the Surrounded by Science project. The goal of the document is to ensure that every project member collects, handles, and stores data following the rules of the project and that the project follows a procedure using a template (of the University of Twente) that is accepted by the European Union. It presents how the project team will manage the data during the project and how the project team intends to share the collected data after the end of the project.

In the project, a Digital Toolbox will be developed through which data will be collected. The Digital Toolbox contains two categories of data. The first category is functional data, which is data that the Digital Toolbox needs in order to function. It involves: (a) project member accounts, (b) science organisations, (c) artefacts, (d) interactive/research elements, (e) experimental elements, (f) project member access data. The second category is research data, which is data that is collected through the Digital Toolbox for research purposes. It involves: (a) participant accounts, (b) participant action data, (c) participant state data, (d) science organisation self-assessment answers, and (e) data collected during experiments. In addition to data collected through the Digital Toolbox, there will also be situations in which data will be collected outside the Digital Toolbox. This data involves data collected from: (a) observational studies, (b) interviews with visitors of science organisations, and (c) interviews with stakeholders. For all types of data, a description is given in this deliverable, including what the data is used for, who owns the data, who can edit the data, and who has read access to the data.

When researchers of the Surrounded by Science project collect data, they will follow the data minimisation principle, so that the amount of personal data to be collected is kept to a minimum. If personal data is collected, participants are asked for their informed consent and the personal data will anonymized. Also, when the users register in the Science Chaser app, they will be asked for their consent.

Data will be stored on certified servers at the University of Twente. It will be shared for open science after the end of the project.

This is the initial version of the data management plan. The project sees the data management plan as a living document and will update it when necessary. In addition to the description of the data management plan in this deliverable, the project has submitted the data management form that is accepted by the European Union to the University of Twente. The approved form has been added to this deliverable as an appendix.

Table of Contents

1 Introduction

This deliverable describes the initial data management plan (DMP) of the Surrounded by Science project. The data management plan is part of WP1, Management and Coordination.

A Data Management Plan (DMP) describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project[1]. It is about collecting and handling the (research) data during and after the end of the project. The following aspects will be described in this deliverable:

- which data will be collected and processed in the project
- how the data minimisation principle is applied when collecting the data
- which data collected by the project falls under the GDPR
- how data will be curated and preserved (including after the end of the project)

After this introduction, Chapter 2 explains how the data will be described in this deliverable. It introduces the terms that are used in the document and the aspects that will be specified for each type of data. Chapter 3 describes the data that will be collected by the Digital Toolbox and Chapter 4 describes the data collected outside the Digital Toolbox. Chapter 5 describes which measures the project will take to keep the majority of the collected data anonymous and which part of the collected data will fall under the GDPR rules. Information on data preservation is given in Chapter 6. The document ends with the conclusions presented in Chapter 7. In addition to these descriptions, the data management form that has been submitted to and approved by the University of Twente is added as an appendix.

The DMP is a living document and will be updated when necessary.

This document is related to deliverable D3.1, which – among others – describes the way the Digital Toolbox handles and stores data, deliverable D8.1, which describes the ethical approval procedures, and deliverable D8.2, which describes the processing of personal data.

[1] Guidelines on FAIR Data Management in Horizon 2020, Version 3.0, 26 July 2016 https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

2 Describing data

The project will collect data through the Digital Toolbox and outside the Digital Toolbox. The types of data that will be collected under both circumstances will be described in Chapter 3 and 4, respectively. Before describing the data that will be collected, we first need to define the terms that are used in this data management plan and we need to specify which aspects of the data will be described.

2.1 Definitions of terms

We define the terms that are used in this data management plan as follows:

Project. The project is the Surrounded by Science project.

Digital Toolbox. The Digital Toolbox is a research instrument to collect data about informal science learning. It consists of four elements: the backend, the Digital Toolbox manager, the Science Chaser and the Science Booster.

Backend. The part of the Digital Toolbox that stores all the data needed for running the Digital Toolbox and that stores the data that is collected by the Digital Toolbox. For more information, see deliverable D3.1.

Digital Toolbox manager. The part of the Digital Toolbox that manages the stored data. It handles the editing of the content presented by the Science Chaser and gives project members access to the collected data. For more information, see deliverable D3.1.

Science Chaser. The main application of the Digital Toolbox, which collects research data about the visitors of science organisations. For more information, see deliverable D3.1.

Science Booster. The application of the Digital Toolbox, that will give recommendations to the science organisations to align the activities they offer with their intended goals. For more information, see deliverable D3.1.

Project member. All the people that are a member of the Surrounded by Science consortium and that work for the Surrounded by Science project.

Participant. A person using the Science Chaser.

User. A person using the Digital Toolbox. That are the project members and the participants.

Visitor. A person visiting an activity of a science organisation. If a visitor is using the Science Chaser during the activity, s/he is also a participant.

Science organisation. An organisation that offers informal science learning activities.

2.2 Aspects of data

For each data type, the following aspects will be specified:

- Description: a short description of the type of data
- Usage: a description of what the data is used for
- Owner: a description of who owns the data
- Edit: a description of who can enter and edit the data
- Read access: a description of who can see the data directly or indirectly.

In general, only the owner of each type of data may edit the data and delete it or may request it to be deleted. One exception to this is that the administrators of the Digital Toolbox will have read, write, and delete access to all data. This can be through the Digital Toolbox manager or through direct access to the locations where the data is stored. The following staff members of the University of Twente will have administrator access: Shant Derounian and Jakob Sikken.

Administrator access will only be used when administrative privileges are needed, for instance when creating accounts for project members. All access through the Digital Toolbox manager will be logged, including access from the administrators.

3 Data collected through the Digital Toolbox

This chapter describes the types of data present in the Digital Toolbox. There are two categories of data: (a) functional data and (b) research data.

3.1 Functional data

The Digital Toolbox needs data to function. It involves data about: (a) project member accounts, (b) science organisations, (c) artefacts, (d) interactive/research elements, (e) experimental elements and (f) project member access data.

3.1.1 Project member accounts

Description. Project members can only use the Digital Toolbox when they have an account. An account consists of a user name/password combination for authentication and access information for authorization. Before project member can use the Digital Toolbox, they will have to give consent by agreeing to the terms and conditions.

Usage. Accounts will be used for accessing the Digital Toolbox manager.

Owner. The project member is the owner of the account.

Edit. Only the project member may edit the properties of the account. When project members request to delete their account, the account will be permanently disabled and the data (described in the following sections) generated by that account is kept. The data of the project member (i.e., the artefacts, content elements, etc. entered earlier by this person) is essential for the Digital Toolbox to function, including its identity. The procedure that is followed for deleting an account will be described in the terms and conditions.

Read access. The Digital Toolbox will read account information for authentication and authorization to give access to the Digital Toolbox.

3.1.2 Science organisations

Description. The science organisation is represented by the project members working for that science organisation. A science organisation has a description and is responsible for all its artefacts.

Usage. A science organisation is shown in the Digital Toolbox as a section with its description and all its artefacts.

Owner. The science organisation is the owner.

Edit. Only members of the science organisation may edit its description and its artefacts.

Read access. Users of the Digital Toolbox can see the descriptions of all science organisations and their artefacts.

3.1.3 Artefacts

Description. An artefact represents a part of an activity offered by a science organisation. It has a description and contains a number of content elements.

Usage. The participants are visiting the artefacts in the real world and can interact with them through their content elements in the Science Chaser.

Owner. The science organisation is the owner of the artefacts of its own organisation.

Edit. Only members of the science organisation may edit its artefacts and content elements.

Read access. Users of the Digital Toolbox can see all artefacts (and their content elements) of all science organisations.

3.1.4 Interactive/research elements

Description. The interactive/research elements are shown in the Science Chaser. Possible examples of elements are questionnaires and games. An element can be tailor-made for a single artefact or be more universal and used for multiple artefacts.

Usage. The elements are used to collect information from the participants. Elements can also be used to motivate the participants, especially game elements will be used for this.

Owner. The project is the owner.

Edit. Project members can edit the interactive/research elements.

Read access. The participants can interact with the elements in the Science Chaser.

3.1.5 Experimental elements

Description. Within the project, a number of experiments will be conducted. For every experiment, there will be a separate experiment element.

Usage. An experiment will be used as a selection criterium to show or hide artefacts, content elements, and interactive/research elements in the Science Chaser. In this way, the Science Chaser will be specifically configured for each experiment.

Owner. The project is the owner.

Edit. Project members can edit experimental elements.

Read access. A participant can see at which experiment s/he is currently participating.

3.1.6 Project member access data

Description. Each relevant interaction of a project member with the Digital Toolbox manager will be stored as access data.

Usage. The access data can be used to analyse technical problems with or possible abuse of the Digital Toolbox member. It can also be used to analyse the usage of the Digital Toolbox manager.

Owner. The project member is the owner of his or her own project member access data.

Edit. Only new access data can be added, existing access data cannot be edited or deleted.

Read access. Only administrators can download the access data. This download action is logged as an access.

3.2 Research data

The Digital Toolbox will also collect research data. It will collect the following types of data: (a) participant accounts, (b) participant action data, (c) participant state data, (d) science organisation self-assessment answers, and (e) data collected during experiments.

3.2.1 Participant accounts

Description. The participants can only use the Science Chaser with an account. An account consists of a user name/password combination for authentication and access information for authorization. Participants' accounts will not have information that can identify the participant, like date of birth, first or last name, address, phone number, email address or photo.

Following the data minimisation principle, a participant's account will only have properties that are needed for the research goals and correct functioning of the Digital Toolbox. The following properties will be asked: age range (e.g., <12, 12-16, >16), level (e.g., beginner, intermediate and advanced), gender, preferred language, and country of residence. These will be asked for as general as possible and as specific as needed (for example using age ranges instead of age in years or date of birth).

Before a participant can use the Science Chaser, they will have to give consent by agreeing to the terms and conditions. The terms and conditions will be in their preferred language.

Usage. Participant accounts will be used for accessing the Science Chaser and as an identifier for analysing the collected data.

Owner. The user is the owner of the account.

Edit. Only the participants may edit the properties of the account, although they cannot edit their assigned user name, to prevent that they enter their real name. When participants request to delete their account, the account will be permanently disabled and the research data generated by that account is kept. As the identity of the participant is unknown, its data is already anonymous. The procedure that is followed for deleting an account will be described in the terms and conditions.

Read access. The Science Chaser will read account information for authentication and authorization to give access to the Digital Toolbox. The participant's account properties will be shown to the project members, for example when included in the export of collected data for analyses.

3.2.2 Participant action data

Description. Each relevant interaction of a participant with the Science Chaser will be stored as action data.

Usage. The project will analyse the participant action data (the interactions) to achieve its research goals.

Owner. The participant is the owner of his or her own participant action data.

Edit. Only new action data can be added, existing access data cannot be edited or deleted.

Read access. The project members can download all access data. This download action is logged as a project member access.

3.2.3 Participant state data

Description. The state of all interactive/research elements in the Science Chaser, which the user can directly or indirectly change, form the state data. These can be answers to a questionnaire, the current state of a game, or achieved scores.

Usage. The participant state data is used to store the state of the Science Chaser for the user. The project can also use the state data to achieve its research goals.

Owner. The participant is the owner of his or her own participant state data.

Edit. The participant can directly (give or change the answers of a questionnaire) or indirectly (increase scores by reporting iSTEM activities) change the state.

Read access. The project members can download all state data. This download action is logged as a project member access.

3.2.3.1 Sharing of reported participant activities

The participants can report iSTEM activities that they engaged in. The reported iSTEM activities will not be directly shared with other participants, to prevent the sharing of illegal and possibly harmful iSTEM activities. Instead, the science organisation project members can create a new artefact from a reported iSTEM activity. Additional background information and explanations can be added. The newly created artefact will appear in the list of suggested artefacts of the science organisation.

3.2.4 Science organisation self-assessment answers

Description. The science organisation will answer a number of self-assessment questions provided by the Science Booster.

Usage. The Science Booster uses the given answers to provide recommendations to the science organisation.

Owner. The science organisation is the owner.

Edit. Only members of the science organisation may answer the self-assessment questions for a science organisation.

Read access. The project members can see the self-assessment questions and the answers and use them for research purposes of the project.

3.2.5 Data collected during experiments

The project will also perform experiments with groups of participants, such as children in a school class or visitors of a science organisation. Besides the information that is collected by default by the Digital Toolbox, additional information will be collected, for example, by conducting pre- and post-tests. This additional information will also be collected through the Digital Toolbox. The experiment itself will be done inside the Digital Toolbox. As a result, there is no special experiment data generated and stored. The data collected during experiments has been described in the earlier sections of the current Section 3.2.

4 Data collected outside the Digital Toolbox

This chapter describes the data that is collected outside the Digital Toolbox. It involves data about: (a) observational studies, (b) interviews with visitors, and (c) interviews with stakeholders.

4.1 Observational studies

The project will perform a number of observational studies. There will be two ways to observe the visitors of an activity: physical observations and website usage.

4.1.1 Physical observations

Description. During a physical observation, the project team will collect data about the interaction of the visitors with a specific science activity, such as counting the number of visitors of a specific science organisation, the duration of these visits, the order in which artefacts are visited, the time of interacting with these artefacts, etc.

The project will not collect any identifiable information of the visitors.

Usage. The collected data will be used for the research in the project.

Owner. The project is the owner.

Edit. During and shortly after the physical observational study, the data will be entered in one or more digital documents. It will not be changed later.

Read access. The digital documents will be stored on a server at the University of Twente (UT). All project members can access the documents.

4.1.2 Website usage

Description. The project will ask science organisations for the usage data of informal science activities on their website. The project might ask the science organisation to add a few questions to the science activities. The expected data is the number of visitors, the duration of their visits, the elements that have been viewed, and optionally a list of answers given to a small survey.

The project will not collect any identifiable information of the visitors.

Usage. The collected data will be used for the research in the project.

Owner. The science organisation is the owner.

Edit. During and shortly after the website usage study, the data will be entered in one or more digital documents. It will not be changed later.

Read access. The digital documents will be stored on a server at the UT. All project members can access the documents.

4.2 Interviews with visitors of a science organisation

Description: In order to get more information from the visitors, a subset will be interviewed. An interview can be held in different ways: using an online form, on paper, or face-to-face. During face-to-face interviews, an audio recording may be made in some occasions.

The project will not ask any identifiable information of the visitors.

Usage. The collected data will be used for the research in the project.

Owner. The project is the owner.

Edit. Shortly after the interviews, the collected answers will be entered in a digital document. It will not be changed later. Afterwards, the paper forms and audio recordings will be destroyed or deleted, this will be done no later than 6 months after the interview.

Read access. The digital documents will be stored on a server at the UT. All project members can access the documents. Access to the audio recordings or paper forms will be limited to a minimum number of project members, for example the interviewer and the one making the transcription.

4.3 Interviews with stakeholders

Description. The project will conduct interviews with stakeholders. The interviews will be held face-to-face, via telephone, or in an online meeting. An audio recording of the interview will be made.

Usage. The collected data will be used for the research in the project.

Owner. The project is the owner.

Edit. Shortly after the interviews, the results will be entered in a digital document. This might include the names of the interviewed stakeholders. At a later time, clarifications might be added.

Read access. The digital documents will be stored on a server at the UT. All project members can access the documents.

5 Data privacy (GDPR)

The project will follow the spirit of the GDPR to collect only the minimal needed data (i.e., data minimisation principle), especially, in terms of personal data.

5.1 Data collected through the Digital Toolbox

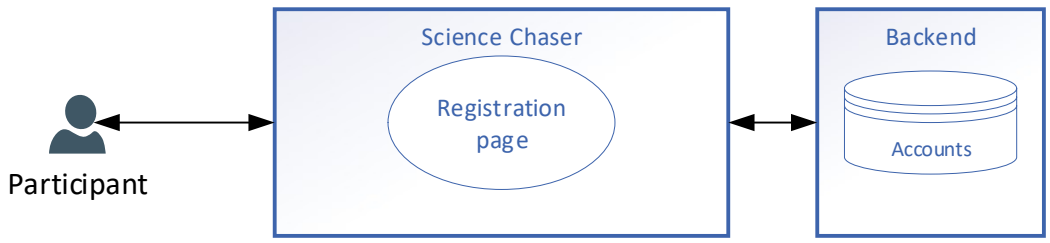
5.1.1 Creating an account

The research goals can be reached without knowing the identity of the participants, therefore no identifiable information of the participants will be collected. During the registration process of the participants, no email address or a phone number will be asked. Instead, the participant will get an anonymous “user name” and s/he has to make up a password him- or herself.

Below are the steps for a participant to register with the Science Chaser and Figure 1 shows the involved parts of the Digital Toolbox (steps and figure are copied from Section 5.2.4 from D3.1):

1. A participant goes to the public website of the project
2. The participant downloads/installs the Science Chaser
3. The participant opens the Science Chaser
4. The participant selects the desired language
5. The Science Chaser shows the consent form with the information about which data will be collected and how it will be processed
6. If the participant gives consent, s/he clicks on the agree button. Clicking the disagree button means that the registration ends at this point and the participant cannot use the Science Chaser
7. The Science Chaser shows a registration screen
8. The participant fills in a password that s/he made up him- or herself, the preferred language, age category, level, gender, and country of residence.
9. This information is sent to the backend
10. The backend creates an account with a random user name
11. The backend sends the created username to the Science Chaser
12. The Science Chaser shows the created user name
13. The Science Chaser strongly recommends the participant to remember the user name and password for future login
14. The Science Chaser opens the account property screen

Figure 1. *Participant’s account registration*



5.1.2 Participating in an experiment

There are two possible situations in which participants are asked to participate in an experiment. In the first situation, visitors of a science activity are asked to participate in an experiment. This is done by sending an invitation for the experiment to all participants that fulfil certain criteria, for example participants that selected a certain science organisation or activity in a certain timeframe in the Science Chaser. The Science Chaser will show the consent form with the information about

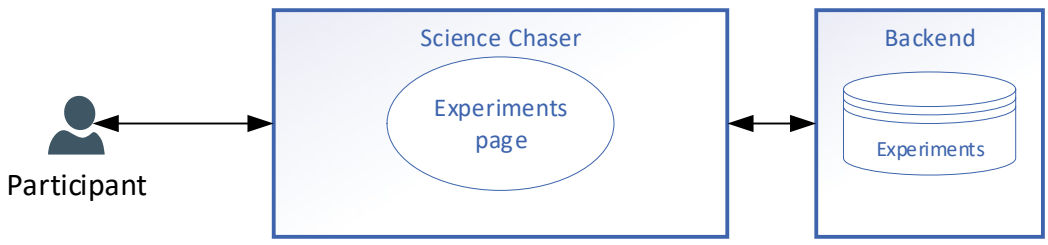
which data will be collected and how it will be processed. The participant has to give his/her consent in order to join the experiment.

In the second situation, a group of people (e.g., a school class) is asked to participate in an experiment. This is done by asking them to join an experiment in the Science Chaser. They will receive an experiment code from the researcher conducting the experiment. Only people with the correct code can join the experiment. For each participant, the researcher will check whether the consent that is necessary has been given. In the Science Chaser, the participant will also see the consent form with the information about the data that will be collected and how it will be processed and s/he has to give consent in order to join the experiment. The project will know which school class and which participants (i.e., their randomly generated user name) joined the experiment, but the project cannot identify which person is behind which participant account.

Below are the steps for a participant to join an experiment and Figure 2 shows the involved parts of the Digital Toolbox (steps and figure are copied from Section 5.2.11 from D3.1):

1. A participant opens the Science Chaser
2. There are two ways to join an experiment
    a. Invited by the Science Chaser
        i. The participant selects a science organisation
        ii. The Science Chaser shows the invitation to join an experiment
    b. Invited by a researcher
        i. A researcher checks if consent is given
        ii. The participant enters a code received from the researcher
3. The Science Chaser will show (again) the consent form with the information about which data is collected and how it will be processed
    a. If a participant does not give consent, the participant can continue their visit to the science activity as normal, but does not participate in the experiment. This is the end of step description
    b. If a participant gives consent, the steps continue
4. A flag will be set in the participant's account that s/he is participating in the experiment
5. The Science Chaser will now show additional questionnaires (e.g., pre- and post-tests) to the participant during the visit
6. An experiment flag will be added to the logging of the relevant actions of the participant
7. When the participant selects "leave science organisation", the experiment flag will be removed from the participant's account
8. If the participant forgets to select "leave science organisation", the Science Chaser will remove the experiment flag when the experiment has ended

Figure 2. *Participants joining an experiment*



The project will not collect additional information of the subjects of the experiment, besides what the Digital Toolbox already has.

An experiment will use the same Digital Toolbox as regular users, and will not create special "experiment" participant accounts. The advantage of this is that the project can connect the

participant's data of the experiment to the participant's data that has been or will be collected by the Science Chaser before and after the experiment.

In future experiments, there might be research elements that involve personal data. An example is when participants are allowed to answer questions with voice recording or photos. Voice recording involves personal data and a photo might contain personal data. In this case, the researcher will explicit check if the needed consent is given. The original voice recordings and photos will be temporarily stored and are not accessible for data analyses. The voice recordings and photos are only accessible as single elements for transcribing and anonymization. The voice recordings will be replaced as soon as possible by the transcribed text and the photos will be made anonymous, for instance a description of the photo will be given or identifiable parts will be removed or blurred. After the voice recordings are replaced by their transcription and photos are anonymized, the original voice recordings and photos are deleted. The anonymized content can be used for data analyses.

The Digital Toolbox has no identifiable information of the participant's account and the additional information collected during experiments does not contain identifiable information. As a result, the collected data from the participants is anonymous and does not fall under the GDPR rules. If in the future, the project decides to collect personal data in the Science Chaser (e.g., voice recording or photo upload), these elements will fall under the GDPR rules.

5.2 Data collected outside the Digital Toolbox

5.2.1 Visitors of science organisations

The data that will be collected outside the Digital Toolbox (during observations and interviews) from visitors of science organisations, will not contain any identifiable information about the visitors. Data is anonymous and does not fall under the GDPR rules. At the start of an interview, it will be explained to an interviewee what will be done with the collected data and the interviewee will be asked for consent.

If the interviews are paper-based or using audio recording, the interviewee could be identified. The answers to the questions will be entered in a digital document and the papers and audio recordings will be destroyed or deleted. This will done as soon as possible. Only a minimum number of project members will have access to the papers and audio recordings.

Even if the interviewee would be identified, it is not possible to link the identity to their participant account (if the interviewee has one). As long as any paper based form and audio recordings of the interview are there, they fall under the GDPR rules.

5.2.2 Stakeholders

The names and email addresses of the stakeholders, whom the project team will interview, are stored by the project. The interviews will be about the professional observations, ideas, and opinions of the stakeholders on informal learning. At the beginning of the interview, the interviewer will explain what the project will do with the interview data. The interviewee will be asked to sign a consent form.

6 Data preservation

6.1 Storage location and compliance with security standards

6.1.1 Location

The collected data will be stored on one or more servers at the UT, which are ISO 27001 as well as NEN 7510 certified. The servers are located in the two data centres on the campus of the UT.

6.1.2 Maintenance

The used servers will be managed by the central ICT department of the UT. The servers will be continuously monitored, and in case of problems, mails will be sent to the maintenance staff of the central ICT department and the server administrators within the project.

6.1.3 Privileged access

Two staff members (Shant Derounian and Jakob Sikken) of the University of Twente will have full access to the Digital Toolbox and everything in it. These staff members are under UT contract and have to comply with the UT data management policy guaranteeing confidentiality. They lose their access if they leave the UT.

6.1.4 Backup and restore

The project will make daily backups of the collected data on the servers of the UT. The backups will be stored at the UT and kept for 28 days. As the backups are made daily, the maximum time of data loss is 24 hours. In case of an incident, the data will be restored within 48 hours.

The backups are intended to be used for a full restore of the Digital Toolbox in case of a disaster with the storage of the Digital Toolbox.

It might be possible to restore parts of the content of the Digital Toolbox, for example in the case of accidental removal of a science organisation and its content. However, this will cost significant effort. The project does not plan to restore parts of the Digital Toolbox.

6.1.5 Versioning management of changes in data

The source code of the Digital Toolbox will be stored at GitHub (https://github.com/). The Digital Toolbox will have an option to export the collected data so it can be analysed. Possible internal changes in the Digital Toolbox can be hidden by keeping the same export format.

Each observational study and experiment will have its own directory. Inside the directory, there will be the following sub-directories: methods, raw data, data ready for analyses, and results.

The recommended naming convention of data files is:

SbS_<study name>_yyyymmdd_<data collection method>_<file name>_v<nr>.<file format>
For example: SbS_ant-exhibition_20220222_observation_durations_v1.0.xlsx

6.2 Analysing the data

The project members can download the collected data of participants, either from the Digital Toolbox or from the data files that have been collected outside the Digital Toolbox. The data will be analysed on the computers of the project members. This will be done according to the security and privacy policies of their organisations.

6.3 Main risk to data (security)

6.3.1 Access management

Only project members will have access to the Digital Toolbox manager. The project members will have to logon with a username/password combination and where possible two factor authentication will be used.

The participants will only work with the Science Chaser. During the registration process in the Science Chaser, an account will be created automatically.

All network traffic between the Digital Toolbox and the outside world (project members and participants) will use up to date standards and secure protocols (HTTPS/SSL).

The data collected outside the Digital Toolbox will be stored at the University of Twente. Project members need to login (using a username/password combination and two factor authentication) to access it.

6.3.2 Auditing of data usage

All relevant interactions with the Digital Toolbox manager (by project members) and the Science Chaser (by the participants) will logged. The Science Chaser logs are used for research purposes of the project.

All the logging can be used for auditing the usage of the Digital Toolbox.

6.4 Open science and data sharing

The project has just started, there is no data yet to share. The collected data will be shared after the project has ended. At the time of writing this deliverable, the data will be shared in about three years. The expected data sharing details are mentioned in the approved UT data management plan, which can be found in Appendix I.

6.4.1 Anonymization

The data about the participants is collected anonymously or is already anonymized. As a result, there is no need for additional anonymizing.

6.4.2 Moment of data sharing

The collected data will be shared after the project has ended. The project members will first get the opportunity to analyse the data and publish the results in one or more papers. No later than 9 months after the project has ended, the collected data will be shared.

6.4.3 Location of data sharing

The project will store the collected data in the publicly accessible repository DANS (https://easy.dans.knaw.nl/ui/home).

6.4.4 Formats of data

The Digital Toolbox will have an option to export the collected data, so it can be analysed. The data will be exported in an appropriate format for analysing. This format will be suitable for sharing the data. The expected format will be a combination of CVS, JSON, and PDF.

6.4.5 Metadata

The generic Dublin Core metadata standard will be used. For better understanding of the data, data files will be accompanied by a copy of the publication and detailed documentation, such as methods, protocols, variable explanation, R scripts/SPSS syntax, and a README file that describes the content of each data file and the relationship among these data files. The shared data will have all the metadata needed to make the data useful for third parties.

6.4.6 License

The data will be made available under the Public Domain Dedication (CC0) license (https://creativecommons.org/publicdomain/zero/1.0/).

7 Conclusions

In this deliverable, a description has been given of the data management plan (DMP) of the Surrounded by Science project. It is an important document, as it provides guidelines for all consortium members on how the project will collect data and how data should be handled and stored. The document describes what kind of data is collected. It also indicates that we will follow the data minimisation principle and it describes how the amount of personal data to be collected is kept to a minimum. The users of the Science Chaser (i.e., the participants) will stay anonymous; the project will not collect any identifiable information about them.

This DMP is written at the beginning of the project and as such reflects the ideas about data of the Surrounded by Science consortium at that time. The project sees the DMP as a living document. When the actual data collection is prepared and done, the DMP will be updated where necessary.

8 List of appendices

Appendix I DMP and GDPR Registration UT

Appendix I

DMP and GDPR Registration UT

DMP and/or GDPR Registration

Introduction

Version: UT-DMP-GDPR_v3-5

Welcome to the UT tool for writing your data management plan (DMP) and the notification of processing of personal data in research in compliance with the General Data Protection Regulation (GDPR registration).

This DMP form has a generally accepted structure which complies to the policy of funders like NWO and ZonMw. The EU allows you to deliver a DMP based on this form as well.

When filling in the answers, please check data policies and guidelines of your research group, department, or faculty and/or the UT research data management policy. Also the website of these organization units may contain relevant information.

This DMP form can or must be reviewed. When your draft version is ready for review press "Save and Review" to start/continue this review process. An e-mail message is sent to the Reviewer. The same holds for GDPR registration.

Choose your form

What do you want to do?

- ☐ DMP
- ☑ DMP with GDPR Registration
- ☐ GDPR Registration

*When you will **process personal data*** in terms of the European General Data Protection Regulation (GDPR), choose option 2 or 3, the latter for instance in case you have already a DMP. UT bachelor or master students are not obliged to make a DMP but in case of processing personal data should do a GDPR registration (option 3). More information.*

A workflow of GDPR and research data (appropriate use of personal data in scientific research according to the GDPR) can be found on this website.

** **processing** means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.*
***personal data** means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;*

The DMP-part of this registration has been accepted by the reviewer.

The GDPR-part of this registration has been accepted by the Data Protection Officer (DPO).

General

G1. Short title of the research

Surrounded by Science

G2. Full title of the research

Surrounded by Science: Learning paths towards science proficiency

G3a. Name of the researcher

Sikken, J.

If this is not the right name, please change it (use the same format)

G3b. Email Address of the researcher

j.sikken@utwente.nl

If this is not the right email address, please change it

G4. Name of the supervisor or principal investigator (PI)

Eijsink, T.H.S.

Please fill in as the following examples: Einstein, A. ; Tocqueville, A.C.H. de ; Waals, J.D. van der

G5. Name of research group/department

BMS-IST

G6. Faculty

Behavioural, Management and Social sciences (BMS)

Please, type the abbreviation of your research group in format:
***Faculty-Research Group**; (e.g. EEMCS-HMI, ITC-ESA, BMS-HTSR)*

G7. Description of the research

Every day, people of all ages are engaged in powerful science-related activities outside the classroom that spark interest and foster their understanding of science. However, it is not fully understood how outcomes of such activities relate to the outcomes of traditional formal science education. Surrounded by Science brings together experts in science education research, science centers and museum pedagogues, providers of outreach and informal learning activities, strong users communities, and policy makers to design and develop a systematic assessment methodology that will analyze the impact of out-of-school science activities.
The project team will conduct a series of field studies and use innovative research instruments (Digital Toolbox) to support the data collection from citizens actively participating in science related activities to assess the impact of specific out-of-school activities. The project will integrate the collected information in matrices to advise science organizations in the design of their activities. Furthermore, the project will propose an accreditation scheme for science education providers and offer guidance on their effective integration to (formal) science education policies. In total, more than 10.000 participants will become part of the assessment, either by interacting with one of the selected activities or by participation in an in-depth assessment of their science proficiency.
Surrounded by Science aims to have a long-lasting impact on the ability of science educators to design more effective, targeted, and meaningful learning activities at all levels. The Digital Toolbox is available to partners within and beyond the project, allowing them to evaluate and set up more effective activities to achieve meaningful impacts on learners? science proficiency. It seeks to enable formal science education to benefit by providing schools with recommendations on how to integrate successful activities to formal learning settings.

G8. Funding body

☐ NWO ☑ EU ☐ ZonMw ☐ Other

In case of NWO, EU and ZonMw, this form is in accordance with their policy. Please note that the ERC Data Management Plan is not a part of the Ethics Review. It is the responsibility of the Principal Investigator to inform the ERCEA Ethics Team of any ethics issues/concerns regarding the collection, processing, sharing and storage of data in relation to the project. The Principal investigator can also be asked to submit an Ethics Data Management Plan (Ethics DMP). For more information you can contact the UT Strategic Business Development office

G9. Grant number

101006349

G10. Partner organization(s) and/or third party(ies)

ELLINOGERMANIKI AGOGI SCHOLI PANAGEA SAVVA AE

EUROPEAN PHYSICAL SOCIETY ASSOCIATION

NUCLIO NUCLEO INTERACTIVO DE ASTRONOMIA ASSOCIACAO

Fondazione IDIS-Città della Scienza

THE LISBON COUNCIL FOR ECONOMIC COMPETITIVENESS ASBL

WEIZMANN INSTITUTE OF SCIENCE

NORGES TEKNISKNATURVITENSKAPELIGE UNIVERSITET NTNU

G11. Project start date

01/10/2021

Please fill in the (provisional) start date of the research project.

G12. Project end date

30/09/2024

Please fill in the (provisional) end date of the research project.

G13. Date written

28/02/2022

Choose the date you have finished this version of the DMP.

G14. Name of the researcher(s), their institution/group and their roles and responsibilities in data management

Name	Institution/group	Role in RDM	RDM Responsibility
Eijsink, T.H.S.	BMS-IST	Project coordinator	Implementing, revising
Gijlers A.H.	BMS-IST		data collection and documentation
Dmoshinskaia, N.	BMS-IST		data collection and documentation
Sikken J.	BMS-IST		writing and updating DMP Custom research software follows DMP rules

In the field 'Name', please fill in as the following examples: Einstein, A. ; Tocqueville, A.C.H. de ; Waals, J.D. van der
Start with who is responsible for implementing the DMP, and for ensuring it is reviewed and, if necessary, revised. Furthermore, indicate for each specific part of the DMP (e.g. data collection, documentation, ...) who is responsible. In case of a collaborative research project, think of the RDM coordination role and responsibility. Check both project organization and research group policies.

G15. Laws, policies, contracts and agreements to comply with

- [x] RDM policy UT / faculty / research group
- [x] General Data Protection Regulation (GDPR)
- [ ] Medical Research Involving Human Subjects Act (WMO)
- [x] Contract (funding bodies, partners, third parties)
- [ ] Other

Please, specify about contract (funding bodies, partners, third parties) to comply with
EU

In principle the handling of data in your research must comply with RDM policies of the UT, faculty and group. However, additional options may apply. When you process personal data your research must comply with the GDPR. Third-party contracts may deal with ownership, responsibilities and requirements about the handling of research data.

G16. Estimated total costs (€) involved in data management

€ 0,00

You can not enter a value here, as this will be automatically filled in as a total of the costs you add in the following sections.

G17. Name of reviewer

Qian Zhang (BMS)

Date of review

01/03/2022

The person in your faculty who can review your DMP.

Data collection

The answers to the following questions should be entered in the table below. Make separate items for each type of data you collect or generate, also specify personal data separately, in order to link them to the GDPR registration section.

1. **Add for each set of data you will collect or generate a descriptive title and the type of data. Think of observational data, experiment data, simulation data and/or derived or compiled data. Also mention materials, such as lab notebooks, field diaries, informed consent, or algorithms, scripts, etc. Add physical data or materials, like samples, as well. Make separate items for personal data.** *Indicate whether it concerns secondary data (pre-existing data collected or generated by other people or organizations)*
2. **Which form will these types of data have (e.g. text, numbers, tabular data, survey data, models, software, audio, video, physical samples)?**
3. **Which file format will the types of data have (e.g. pdf, xls, doc, txt, rdf)?**
4. **Which software or tools are needed to create, process and/or visualize these types of data?**
5. **Do these types of data contain personal data in terms of the European General Data Protection Regulation (GDPR)?** *If yes, tick the checkbox and register the processing of this personal data after finishing the DMP sections.*

DC1. Descriptive title / type of data	DC2. Form	DC3. File format	DC4. Software/tools	DC5. Personal data
Recordings of interviews with stakeholders with informed consent	audio	mp3	media player	☑ Yes ☐ No
Interviews with stakeholders with informed consent	text	doc	ms word	☑ Yes ☐ No
Recordings of interviews with participants with informed consent	audio	mp3	media player	☑ Yes ☐ No
Interviews with participants with informed consent	text	doc	ms word	☐ Yes ☑ No
Observation of participants	numbers	xls	ms excel	☐ Yes ☑ No
Actions and answer of participants with informed consent	text and numbers	json	text viewer	☐ Yes ☑ No
Analyses syntax	text	text	ms word	☐ Yes ☑ No
Informed consent	text	text	ms word	☑ Yes ☐ No

At some point in the research will personal data be anonymized or pseudonymized?

☑ Yes ☐ No

Anonymization of personal data means de-identification that is not reversible: once personal data has been stripped of identifying data, it is no longer possible to trace back to natural persons. Pseudonymization means that personal data are replaced by a key, making it possible to trace back to natural persons.

DC6. In case you use secondary data which source will be used?

Secondary data can be used from very different sources, such as data available in your own research group, from databases managed and offered by (inter)national institutes, e.g. statistical offices; commercial parties or clinical data from hospitals.

DC7. Is copyright on data owned or claimed by a third party?

☐ Yes ☑ No

This question is about rights and control regarding the research data. Although legally incorrect, this is often referred to as data ownership. In case of primary data, intellectual-property rights ('database right') is vested in the University of Twente (see UT research data policy, section 4).

DC8. What will be the estimated total costs (€) involved in the collection, generation and/or use of data?

€ 0,00

Think of costs for acquiring, processing or analyzing the data or for getting informed consent. Use cost estimations which possibly are available in the research project budget. For more information, see Guide Research Data Management and Costs.

Data documentation

DD1. What standards will be used for metadata? If there is no discipline-specific metadata standard available for the data, you could use a general metadata standard (e.g. Dublin Core), and list additional metadata that will be used.

Dublin Core

Data documentation should always contain basic details, called metadata, to make the data findable, including who created or contributed to the data, its title, date of creation and under what conditions it can be accessed. Metadata standards are for instance Dublin Core, DDI, TEI, EML, MARC or CMDI. Data repositories, e.g. 4TU.ResearchData or DANS, use metadata standards as well.

DD2. How will the data (primary and/or secondary) be documented during your research and for long-term preservation?

For all the data in the digital toolbox, detailed TypeScript interfaces will be defined.
For the interviews a list of questions will be prepared.
For the observational studies an observation template will be created.

Data documentation should also include details on the methodology used, analytical and procedural information, definitions of variables, vocabularies, units of measurement, any assumptions made, and the format and file type of the data. Consider how you will capture this information and where it will be recorded, for example in a database with links to each item, a 'readme' text file, file headers, code books, or lab notebooks. In case of secondary data, explain how data provenance will be documented.

DD3a. What directory-naming convention will be used?

Each observational study and experiment will have its own directory,
Inside each study/experiment directory will be the following sub directories:
- methods
- raw data
- data ready for analyses
- results

There is no single recommended way to name your folders/directories, but you should name your files consistently. Check the file-naming system with your colleagues in the group. More practical information can be found here.

DD3b. What file-naming convention will be used?

SbS_<study name>_yyyymmdd_<data collection method>_<file name>_v<nr>.<file format>

There is no single recommended way to name your files, but you should name your files consistently. Check the file-naming system with your colleagues in the group. More practical information can be found here.

DD4. How will you handle version control (tools, software, procedures) to maintain all changes that are made to the data?

The software will be stored at GitHub (https://github.com/).
Each data file will have its own date and version in the file name.

Version-control mechanisms should be established and documented before any data is collected or generated. Read more about version control and specific tools.

DD5. What will be the estimated total costs (€) involved in the documentation of data?

€ 0,00

In case of labour intensive metadata creation and documenting data files, hiring of personnel may be needed. Use cost estimations which possibly are available in the research project budget. For more information, see Guide Research Data Management and Costs.

Data storage

Use the table below to answer the the following questions for each type of data you collect or generate. If you want to add or change an item, please return to the table Data Collection.

1. **What storage medium will you use for the master files of the data?**
2. **In case of non-UT central storage media, what will be the backup frequency and location?**

It is UT policy to store the original files (master files) from which you make work copies of the research data on UT network file servers, such as the 'Home directory' or the 'Project and organization directory'. Data files on UT network servers are stored in the UT data centre and backed up daily. Have a look at the UT options for storing your research and the backup procedure. For questions and more information, contact the ICT Account manager in your faculty.
If you are, by third-party agreement, bound to store the master files on another medium, be aware of adequate backup frequency and location. Storing master files of data on laptops, stand-alone hard drives or portable storage devices such as USB-sticks, is not in compliance with the UT data policy.

DS1. Storage Medium

Type of data	DS1. Storage Medium	DS2. Backup Frequency	Backup location
Recordings of interviews with stakeholders with informed consent	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre
Interviews with stakeholders with informed consent	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre
Recordings of interviews with participants with informed consent	☐ UT Network Storage ☑ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage* Please specify other storage: Project partner	Daily	Project partner
Interviews with participants with informed consent	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre
Observation of participants	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre
Actions and answer of participants with informed consent	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre
Analyses syntax	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre
Informed consent	☑ UT Network Storage ☐ Non UT Network Storage *Cloud services such as Surfdrive or OneDrive (with @utwente.nl login) are non UT network storage*	Daily	UT Data Centre

What will be the estimated total volume of data stored on UT network servers?

Amount of estimated total volume	Units of total volume
10-100	GB (GigaByte)

DS3. In case of storing master files (also) on other media than the UT network file servers, what are the reasons of this?

Only the interviewer, working for a project partner, has access to the recording of interviews with visitors. After the recordings are transcribed, they will be deleted.

If (master files of) data must be (also) stored on servers of external parties, please refer to the contract or agreement.

DS4. If other storage media for work copies of the data files will be used, specify this here.

Although not recommended, copies of data files can be kept on remote, cloud and/or portable storage. Please, keep in mind confidentiality and security requirements.

DS5. What are the estimated total costs (€) for storage of the data, both on UT network servers and other locations?

€ 0,00

When you need to store more than 10 GB of data on UT network servers, ask the ICT Account manager in your faculty for information about costs.

Data security

DSec1. In case of an incident, explain how the data will be recovered.

As the backups are made daily, the maximum time of data loss is 24 hours. In case of an incident, the data will be restored within 48 hours.

Refer to data-recovery procedures of your backup services or facilities, such as how long it will take before you can use the data again, and the time between the incident and the last valid backup. The UT backup service for employees gives you more information. When applicable, refer to specific security measures or protocols.

DSec2. Do (part of) the collected or generated *NON-personal* data have a confidential character?

☐ Yes ☑ No

Data may be confidential, due to political sensitivity, trade interests (including patentability) or public security

DSec3. During the research, who controls access to the data?

The principle investigator

For instance, the researcher, the principal Investigator, the supervisor or head of the research group. Sometimes, institutions (e.g., the lab, the university or even the funder) should control access to the data.

DSec6. After the research, who controls access to the data?

The University of Twente, department of Instructional Technology

In contrast to during the research, data-access control after the research is often managed differently. It could remain with the researcher, principal investigator or the supervisor, but it is better to choose for a more persistent solution, such as the (head of the) research group, or institute.

DSec9. What will be the estimated total costs (€) involved in data security?

€ 0,00

In case of labour intensive metadata creation and documenting data files, hiring of personnel may be needed. Use cost estimations which possibly are available in the research project budget. For more information, see Guide Research Data Management and Costs.

Data selection and preservation

In this section you are asked to think about the data preservation. Use Areda, the UT Archive for Research Data, to preserve your data (Areda can be found at https://areda.utwente.nl). The next section (Data availability for reuse) is about depositing and publishing in a trusted data reposory, such as 4TU.ResearchData and DANS.

Use the table below to answer the the following questions for each type of data you collect or generate. If you want to add or change an item, please return to the table Data Collection. In case of a new item, don't forget to add information, also in the table of Data Storage section. Be aware that secondary data should not be added in this table. You can give details about preservation of secondary data in question 5.

1. **Which of the data will be selected for preservation?**
2. **Which file formats will be used for long-term preservation and accessibility?**
3. **After the end date of the project, how many years should the selected data be preserved?**
4. **What must be done with the data after the end of the preservation period?**

When preserving data, use non-proprietary (open) persistent formats, such as pdf, txt, rif, dot, tif, flac, csv or xml. More information about preferred formats you can find at DANS or 4TU. ResearchData. Preserve the data for at least 10 years (in compliance with the UT data policy) unless there are legal or contractual conditions for choosing another preservation period. In some cases, data files have to be destroyed when the data are not needed anymore for the purpose for which it has been processen, mostly when containing sensitive personal data.

Type of data	DSP1. To be preserved	DSP2. Preservation format	DSP3. Preservation period	DSP4. Post preservation action
Recordings of interviews with stakeholders with informed consent	☐			
Interviews with stakeholders with informed consent	☑	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other
Recordings of interviews with participants with informed consent	☐			
Interviews with participants with informed consent	☑	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other
Observation of participants	☑	.csv	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other
Actions and answer of participants with informed consent	☑	.json	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other
Analyses syntax	☑	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other
Informed consent	☑	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other

What will be the estimated total volume of the selected data to be preserved in Areda?

Amount of estimated total volume	Units of total volume
10-100	GB (GigaByte)

DSP5. Are there secondary data to be preserved?

☐ Yes ☑ No

Whether or not you are allowed to preserve secondary data depends on the terms of use. In either case, explain how accessibility to the data you used will be arranged after the project.

DSP6. For the sake of proper verification and reuse, what preparations will be made before preserving the data?

The audio recordings of interviews will be transcribed.

In some cases, data need to be cleaned, processed further, checked, etc. before being preserved. Also digitization or file conversion may be needed. When personal data is used anonymization or pseudonymization is needed.

DSP7. What will be the estimated total costs (€) involved in the preparation of data for preservation?

€ 0,00

Only costs for preparing data for preservation. Costs of preservation of the data is part of the next section, Data Availability for Reuse. Use cost estimations which possibly are available in the research project budget. For more information, see Guide Research Data Management and Costs.

Data availability for reuse

This section is about depositing in data repositories such as 4TU.ResearchData and DANS.

Use the last three columns of the table below to fill in the answers to the following questions for each type of data you collected or generated. If you want to add or change an item, please return to the table Data Collection. In case of a new item, don't forget to add information, also in the table of Data Storage and the Data Selection and Preservation section.

1. **Which repository is appropriate for depositing the selected data?**
2. **Which of the selected data can or must be made publicly available?**
3. **Which license of use will be assigned to the published data?**

The UT research data management policy, faculty data policies and many funders demand depositing in a trusted data repository (having for instance a CoreTrustSeal / Data Seal of Approval), in principle openly accessible. Choose 4TU.ResearchData for beta-sciences or technical-sciences data, DANS for humanities, health sciences, social and behavioral sciences, oral history and spatial sciences. You can use Zenodo in combination with Github for preservation of software code. Costs for preservation can be found on the website of the data repository. Find other trusted data repositories on the CoreTrustSeal website. Be sure that deposited data sets will get a persistent identifier, such as DOI (Digital Object Identifier) for easy retrieval and sustainable linking and citation.
In general, public access to data is preferred, and many funders demand this. Sometimes embargo is needed. In other cases, e.g. for reasons of confidentiality, data cannot be made public. About assigning licenses to data, you can find more information on the 4TU.RD website.

Type of data	Format	Preservation period	Post preservation action	DA1. Repository	DA2. Public	DA3. License
Interviews with stakeholders with informed consent	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other	☐ 4TU.RD ☑ DANS ☐ Other	☑ Yes ☐ Embargo ☐ No	CC0
Interviews with participants with informed consent	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other	☐ 4TU.RD ☑ DANS ☐ Other	☑ Yes ☐ Embargo ☐ No	CC0
Observation of participants	.csv	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other	☐ 4TU.RD ☑ DANS ☐ Other	☑ Yes ☐ Embargo ☐ No	CC0
Actions and answer of participants with informed consent	.json	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other	☐ 4TU.RD ☑ DANS ☐ Other	☑ Yes ☐ Embargo ☐ No	CC0
Analyses syntax	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other	☐ 4TU.RD ☑ DANS ☐ Other	☑ Yes ☐ Embargo ☐ No	CC0
Informed consent	.pdf	> 10 yrs	☐ To be destroyed ☑ To be decided ☐ Other	☐ 4TU.RD ☐ DANS ☑ Other University archive Areda *Other repository*	☐ Yes ☐ Embargo ☑ No	

For those data that are not made publicly available, what are the reasons for this? Will availability of the data be restricted to a certain group or community, and if so, which, why and how?

Both the UT research data management policy and funder policies demand data to be as open as possible. Possible reasons for not making data publicly available are legal restrictions on personal data, intellectual property, political sensitivity, commercial interests (also patenting), and public security.

DA4. When necessary, how will specific software for viewing, analysing or reusing the data be made available?

No specific software is needed as data is stored in open formats

In some cases specific software is needed for viewing or analysing data. It is important that this software is available for verification and/or reuse of the data.

DA5. Will exclusive or restricted use of the data be claimed by third parties?

☐ Yes ☑ No

DA6. What are the estimated total costs (€) for depositing the data?

€ 0,00

Check any cost arrangement between your faculty with a particular data repository. If no specific arrangement apply, check depositing costs at the website of the repository. Use cost estimations which possibly are available in the research project budget. For more information, see Guide Research Data Management and Costs.

GDPR registration of processing personal data

According to the General Data Protection Regulation (GDPR) registration is mandatory when personal data are processed in research. If personal data are handled in your research, you need to answer the questions below. Your research will then be added to the GDPR register of the University of Twente. Click here for more information. If you have any questions, please contact the Privacy Contact Person in your faculty.

A workflow of GDPR and research data (appropriate use of personal data in scientific research according to the GDPR) can be found on this website.

1. Under the responsibility of which organization is the research project carried out?

- [x] University of Twente
- [ ] Other

*In most cases, the University of Twente (UT) will in principle be regarded as the **controller*** for the personal data processing in the context of the investigation. After all, in most cases the researchers will carry out the research in the context of their employment at the UT. Although it is not inconceivable that the foregoing will be different in specific cases, we believe that the basic principle is that the UT qualifies as a controller.*

** controller: the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.*

2. Who is the contact person of the research project?

Who should be contacted when the Data Protection Officer of the UT has a question about the processing of the data? Most likely, this is the researcher or the supervisor, principal investigator or (co)promotor. In any case, it must be an employee of the UT. When the researcher is a graduate student, the contact person is the supervisor.

Name Type the surname. A list of names will appear. Select one of the shown names.	Department/group	Phone number	E-mail address
Eijsink, T.H.S. (BMS-IST)	Faculty of Behavioural, Management and Social Sciences (BMS)	+31534893572	t.h.s.eijsink@utwente.nl

3. Is a third party involved in the processing of the personal data?

- [ ] Yes
- [x] No

A third party is a non-UT organization or person and is involved in tasks such as adding or analyzing data or elaborating interviews, or even just viewing data. Also IT tasks, such as the technical management and hosting of data infrastructure, can be executed by a third party. In the event of a third party, it will be necessary to determine whether this party should be regarded as a processor or controller. As indicated earlier, the controller is the one who determines the purpose for which personal data are processed and by what means the processing takes place. In other words, it is the person who determines what exactly happens. Being a processor only exists if that party has no, or only minimal, control over the processing.
In order to answer the question of whether someone should be regarded as a controller or processor, it is useful to determine why the data processing takes place and who took the initiative to do so.

4. What is the full title of the research project?

Surrounded by Science: Learning paths towards science proficiency

5. Which of the following legal bases apply to the research project?

- [x] consent of data subject(s)
- [ ] execution of an agreement in which the data subject is a party
- [ ] compliance with a legal obligation
- [ ] protecting the vital interests of the data subject or someone else
- [ ] fulfilment of a task of general interest or public authority
- [ ] representing the legitimate interests of the controller or a third party

In the privacy law, the General Data Protection Regulation (in Dutch: AVG), six reasons are stated for processing data. You need a legal basis to be able to process personal data, otherwise the processing of personal data is unlawful! You are not allowed to change the legal basis, so you must choose it in advance and not justify your processing afterwards. If you have any doubts about which legal basis applies to your research, it is strongly recommended that you contact the Privacy Contact Person (PCP) of the relevant faculty or the Data Protection Officer (DPO).

6. What is the purpose of the collection of personal data?

The project may want to contact the stake holders in the field of informal learning again for additional information or clarification.
For the interview with visitors of science organisations, audio recording and paper based forms are ways to collects answers to a number of questions.

It is important to explain why you need the specific personal data you will process. The personal data that will be processed can only be used for the purpose you formulate here. You are not allowed to process personal data that are not directly relevant (data minimization). Be specific, for example only 'research' is not enough.

7. What is the category of the data subject(s)?

Stakeholders in the field of informal learning.
Visitors of science organisations.

Example: patients, employees, students, children.

8. What personal data will be processed?

This information is copied from the table in the section Data Collection. If you need to change or add information, you have to do so in the Data Collection table. ***Do not forget to update the information in the tables in the sections Data Storage, Data Selection and Preservation, and Data Availability for Reuse****.*

The GDPR provides (in Article 4) the following definition of personal data:"Any information relating to an identified or identifiable natural person ("the data subject"); identifiable is a natural person who can be identified, directly or indirectly, in particular by means of an identifier such as a name, an identification number, location data, a online identifier or of one or more elements characteristic of the physical, physiological, genetic, psychological, economic, cultural or social identity of that natural person. "

This means that all information that directly relates to an individual, or that can be traced back to an individual, is personal data. This concerns data such as a person's name, address, place of residence, telephone number and date of birth. But also customer numbers, salary slips, photos, business telephone numbers, fingerprints, DNA material and e-mail addresses are personal data.
To determine whether something is personal data, it is important not only to look at the data as such. Personal data can also arise through the link between different sources. For example, the registration number of a car as such is not personal data. But if an organization has access to the RDW database, then the license plate can be linked to a person - and the information for the organization in question is therefore personal data. This also applies to, for example, IP addresses of websites and location data of mobile phones. Personal data can also be involved when using big data. By linking different data, information can be (directly or indirectly) traceable to an individual. So there are many types of personal data. By law, it must be information about a natural person. Information about organizations, such as legal persons, associations and foundations, is not personal data. Data from deceased people also falls outside the scope of the GDPR.

Recordings of interviews with stakeholders with informed consent

Interviews with stakeholders with informed consent

Recordings of interviews with participants with informed consent

Informed consent

9. What special personal data regarding the mentioned category of data subjects will be processed?

Be aware of the conditions for processing these special personal data, such as having explicit consent from the data subject for processing those personal data for one or more specified purposes. See GDPR art. 9.
Please note when completing this question. In principle, it is prohibited to process such personal data. If the researcher wishes to process special and / or criminal personal data, there must be an exemption under the GDPR. If you are unsure whether or not special and / or criminal personal data are involved, please contact the PCP of the relevant faculty or the DPO. Beware: in the Netherlands, in addition to special and criminal personal data, we also have a separate category of personal data for which special or stricter rules apply for processing: the Citizen Service Number (BSN). After all, the BSN may only be processed if required by law and regulations.

☑ none	☐ sex life or sexual orientation	☐ religious or philosophical beliefs
☐ trade-union membership	☐ racial or ethnic origin	☐ criminal justice data
☐ political opinions	☐ genetics data	☐ health
☐ biometics data for purpose of uniquely identifying a natural person		☐ Citizen Service Number (in Dutch: BSN-nummer)

10. How long will personal data be preserved?

If you need to change the preservation period, you have to do this in the Data Selection and Preservation table. The preservation period must be related to the purpose of the processing. It is not allowed to keep personal data longer than the purpose for which they were processed. After that time, the data must be erased or completely anonymized. When personal data under a different law have a mandatory preservation period, and this is longer than is appropriate for the purpose of the processing, then this mandatory preservation period applies.
In the context of this question, it is important to tell something about the distinction between the anonymization and the pseudonymisation of personal data. In the case of pseudonymisation, personal data within the meaning of the GDPR is still involved, to which the rules from the GDPR apply. "Anonymization" means that the data can no longer be traced back to a data subject in any way. This means that it is not always sufficient to simply remove name and address details. After all, there may still be information that can be indirectly traced to the data subject, and thus personal data within the meaning of the GDPR. In that case, pseudonymisation will therefore be more likely.

Type of data	Preservation period

Interviews with stakeholders with informed consent	> 10 yrs
Informed consent	> 10 yrs

11. After the end of the preservation period, will the personal data be destroyed automatically or manually?

☐ Automatically ☑ Manually

12. With which external parties will the personal data be shared?

Project members

External parties can be the processor, mentioned in question 3, but also parties that use or view the personal data, for example in the form of reports. Statistical and anonymized data are outside the scope of the GDPR and should not be mentioned here.
"Anonymization" means that the data can no longer be traced back to a data subject in any way.

13. Which measures will be taken to protect personal data against data loss, theft, and/or unauthorized use?

In case of other security measures, please specify.

- ☑ data security policy
- ☐ safe of storage of data files
- ☐ burglar alarm
- ☐ physical measures for access security
- ☐ access by biometric characteristics
- ☐ access by card
- ☐ automatic logging of access (including verification procedure)
- ☐ access by password or PIN
- ☐ access by other means
- ☐ verification of authorizations
- ☐ Other: please specify

14. Will any personal data be transferred electronically?

☑ Yes ☐ No

Type of data	By which network
Recordings of interviews with stakeholders with informed consent	☐ No network ☑ UT network ☐ Other network
Interviews with stakeholders with informed consent	☐ No network ☐ UT network ☐ Other network
Recordings of interviews with participants with informed consent	☐ No network ☑ UT network ☐ Other network
Informed consent	☐ No network ☐ UT network ☐ Other network

15. Which type(s) of encryption will be applied?

- ☐ File encryption
- ☑ Storage encryption
- ☑ Transfer encryption
- ☐ No Encryption

You can choose more than one option. In any case, it must be in line with the information written in your Data Security Section, question 5. Read more about encryption on the LISA website.

16. Will personal data, processed in your research, be transferred to one or more non-EU countries?

☑ Yes ☐ No

List of member countries of the EU.
In principle, transfers of personal data to countries outside the European Economic Area (EEA) are prohibited. In such cases, it is advised to contact the PCP of the relevant faculty or the DPO.

Will personal data only be transferred to non-EU countries that have an adequate level of protection?

☑ Yes ☐ No

List of Non-EU countries with a adequate data protection: Andorra, Argentina, Canada, Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland, Uruguay as providing adequate protection.

Please select the NON-EU countries with a adequate level of protection where personal is transfered to.

- [ ] Andorra
- [ ] Argentina
- [ ] Canada
- [ ] Faroe Islands
- [ ] Guernsey
- [x] Israel
- [ ] Isle of Man
- [ ] Japan
- [ ] Jersey
- [ ] New Zealand
- [ ] Switzerland
- [ ] Uruguay

Scrollable list of countries with a adequate level of protection.

17. Is it necessary to carry out a Data Protection Impact Assessment (DPIA) prior to the start of data processing?

- [ ] Yes
- [x] No

Where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. This is particularly likely when you use new technologies. For more information, see the EU website about this topic. For questions, contact the Privacy Contact Person in your faculty.

18. Statement of approval.

The undersigned declares that the information provided in this GDPR registration form is correct.

Name

Sikken J.

Email

j.sikken@utwente.nl

Function

Researcher

Place

Enschede

Date of GDPR registration

28/02/2022

- [x] **I declare that the information in this GDPR registration is correct.**

GDPR Registration reviewed and accepted by Data Protection Officer (DPO)

01/03/2022

To be filled in by the DPO when this GDPR registration has been adapted in accordance with comments. As long as this date is not filled in, comments will remain visible.

*Thank you for registering the processing of personal data (GDPR) in your research. Please press the **Save & Review** button to send it to the UT Data Protection Officer. You will receive an email once it has been reviewed.*
If you have any suggestions or questions, please contact the Privacy Contact Person in your faculty.