Plan Overview

A Data Management Plan created using DMPTuuli

Title: Students with special needs, teacher survey

Creator: Martti Mäkinen

Principal Investigator: Martti Mäkinen

Data Manager: Martti Mäkinen

Affiliation: Hanken School of Economics

Template: Hanken's DMP template

ORCID iD: 0000-0002-6806-4217

Project abstract:
This project charts teacher and student experiences and pedagogical needs and solutions for students with various special needs. The purpose of this project is 1) to collect information about the current state of teachers' knowledge and experience about the attainability of studies, 2) to raise the awareness of survey participants about special needs among students at institutions of higher education, and 3) to survey students with special needs experiences of accessible studies, or the lack thereof.

ID: 18861

Start date: 29-04-2022

End date: 31-12-2023

Last modified: 21-03-2023

Copyright information:

Students with special needs, teacher survey

Checklist of responsibilities

Check and ensure that you have fulfilled all the following responsibilities:

- Writing up this DMP before the data collection is started;
- Making the DMP available (on request) to research participants (i.e. persons whose data is included in the dataset);
- Submitting a copy of the DMP to dpo@hanken.fi before the data collection is started;
- Modifying and updating the DMP in case changes are planned to the data processing activities and re-submitting a copy of the modified DMP to dpo@hanken.fi immediately after the modifications. (Typical modifications include, e.g., the addition of another research team member's name to the list of persons processing the data, and addition of another data variable to the list of variables compiled into the dataset. Please note, though, that the purpose of the dataset and data processing cannot be modified after the collection of the data has started.)

General information

1.1 The study/dataset is for:

- Other. Please specify below:

The surveys and interviews are part of University of Turku development project 2digi2, in which pedagogical solutions and best practices for online teaching are sought and created. The data collection described in this DMP is carried out by a sub-group of 2digi2, Special Needs, researchers from several universities in Finland.

1.2 If you are a BSc/MSc/eMBA student, who is your thesis supervisor?

N/A

1.3 Give a brief description of your research such as its name, research members, funding information and project number. For example, if it is an official research project at Hanken, what is Hanken's project number?

The members of the team for surveys described here are Martti Mäkinen (Hanken), Signe-Anita Lindgren (Aalto University), Jaana Suviniitty (Aalto University), Riikka Halonen (University of Eastern Finland), Lauri Tolkki (University of Eastern Finland), Laura Lehtinen (UniArts), Laura Senni (University of Lapland), Sirpa Heino (University of Vaasa), and Olesia Kullberg (LAB).There is no funding for the project as of yet.

1.4 What is the version of this DMP? Is it an initial, detailed or final version? Also specify here the date when this version of the DMP is written.

A DMP ought to be created *before* you start your research project and be updated as your research project evolves. It is a living document that accompanies the whole research life cycle, even after the active phase of the research project.

Try to avoid overlapping the content of your DMP with that of your research plan. The research plan describes the scientific, analytical and methodological processing of data, whereas the DMP describes the technical and administrative management of data.

This is a detailed version of DMP, published on May 4, 2022 and revised Feb 15, 2023.

1. Data description

1.1 Give a brief description of your data. Answer the questions:

- How *new data* will be collected or produced in the project?
- How *existing data* will be reused?
- What *kinds* of data will be collected, produced or reused?
- In which *file formats* will the data be in?
- And estimate the *data size* if you know.

New data will be collected through Webropol surveys. There is no pre-exisiting data. The surveys collects answers (both closed and open-ended) to questions about experiences with and by students with special needs, and some non-compulsory, demographic data. The data will be stored in Webropol and exported in Excel sheets.

1.2 Answer the question how the consistency and quality of data will be controlled. Careful documentation of the procedures of data collection is the primary measure to ensure the integrity and quality of the data.

The quality and consistency of data will be ensured by constant documentation of the project.

2. Ethical and legal compliance

2.1 Researchers shall bear the responsibility for ethical and moral concerns and decisions involved in the research and during the interaction between the researcher and research participants. Follow *Hanken's ethical guidelines and good data protection practices* **to maintain high ethical standards and comply with relevant legislations.**

What ethical issues are related to your data management, for example, in handling personal data, protecting the identity of participants, gaining consent for data sharing, and/or handling sensitive data? How will you take care of these ethical issues? Read *Hanken's DMP guidance* **on Ethics & privacy (on the right) carefully when writing the DMP for this section 2.1.**

1. **Right to collect, handle, and publish data**
Our study contains personal data that is gathered from persons or individuals. The research team members (data processors/controllers) follow the data protection policy of Hanken School of Economics (https://www.hanken.fi/privacy/) and comply with relevant data protection laws in a transparent, fair, and lawful manner.
The purpose of the dataset and legal basis for data processing is scientific research carried out in the public interest as defined in Finnish Data Protection Act.
In addition, we will ask for informed consent of the research participants (i.e., from whom the data is gathered). Before data collection, we will inform the research participants of what data we are collecting about them and how, ask their permission for this, and ensure that the participation is voluntary.
The informed consent message that participants are requested to read and accept is attached as an annex of this DMP.

2. **Data controllers and data processors:**
Martti Mäkinen is the data controller. The processors are Martti Mäkinen (Hanken), Signe-Anita Lindgren (Aalto University), Jaana Suviniitty (Aalto University), Riikka Halonen (University of Eastern Finland), Lauri Tolkki (University of Eastern Finland), Laura Lehtinen (UniArts), Laura Senni (University of Lapland), Sirpa Heino (University of Vaasa), and Olesya Kullberg (LAB) who conduct transliteration/transcription work and the analysis of data.
A data processing agreement (DPA) will be made with Signe-Anita Lindgren (Aalto University), Jaana Suviniitty (Aalto University), Riikka Halonen (University of Eastern Finland), Lauri Tolkki (University of Eastern Finland), Laura Lehtinen (UniArts), Laura Senni (University of Lapland), Sirpa Heino (University of Vaasa), and Olesya Kullberg (LAB)..

3. **Types of data collected**
To maintain the trust of research participants and to protect their identity, the project team members will ensure that the personal data collected is pseudonymised or anonymised. First, during data collection, the researchers will use general terms to refer to data subjects' participation in project activities. This is to ensure that all personal data will be pseudonymised, and that data subjects are not identifiable without the use of separately stored additional information. In addition, information on the original values and techniques used to create the pseudonyms will be kept organisationally and technically separate from the pseudonymised data.

Before archiving the research data, pseudonymous data will be made anonymous by destroying the separately kept identifying information. The research team make an anonymisation plan which describes the anonymisation measures and evaluates the disclosure risk of data subjects' personal data. The project will save the anonymised data in a non-proprietary format (e.g., plain text) and deposit it into the data storage repositories.

In detail, the following **direct identifiers,** email address and information about the employer, when collected from the respondents will be stored in a separate file, and will be *removed from the dataset in 4 weeks after obtaining the persons' response or data*. Research participants will be able to request to see their personal data, to correct them, or to erase them until the said date, when the data that can directly identify them has been erased from the dataset.
The data collected **does not involve** **sensitive personal data**. Therefore also no data protection impact assessment (DPIA) is needed.
The contact information of research participants/subjects/respondents/informants can be volunteered. In that case it is stored in a separate file which is not linked to (and cannot be linked to) files containing other personal data. The purpose of maintaining the contact information is to be able to contact the research participants if a follow-up interview is organised. The contact information data file will be erased 4 weeks after the project is completed.
Collected personal data include the research subject's contact information and workplace. The processed data is obtained from the data subjects directly on the basis of consent.

5. **Potential archiving of data**
In case the data would be archived for later use after the project has ended, all the above-mentioned data types (direct identifiers and contact information) will be deleted from the dataset before we submit the data for archival.

The project will be given the permission by research participants to share their data in an anonymised form on the project website, in the project deliverables, in journal articles and other publications, and in the data storage repositories.

6. **Transfer of data**
The research team members will rigorously maintain high ethical standards and comply with relevant legislation. The personal data collected will not be transferred outside the European Union. All data transferred between project partners (within the EU) will be restricted to pseudonymized or anonymized data and the transfer will only be made in encrypted form via a secured channel.

2.2 Agreements on data ownership and other intellectual property rights such as secondary data usage copyright permissions and open data licenses need to be concluded before commencing any actual research activities. In the data ownership agreement, describe who owns the data, and whether and what rights will be transferred. Agreements about authorship need to be done before the beginning of the project.**

How will you manage the rights of the data you use, produce, and share?

The principal investigator is responsible for concluding contracts on authorship, data ownership, data sharing, and user rights, which will be agreed on with all researchers prior to the start of actual research. The data ownership agreement describes who owns the data, and whether and what rights will be transferred. Copyright and intellectual property rights will also be secured before any data is made public.
When reusing the secondary data gathered from open sources, good practices for the attribution of authorship and data citation will be followed, and all legal restrictions such as copyright permissions and license terms on its use observed.
The principal investigator is responsible for concluding contracts on authorship, data ownership, data sharing, and user rights, which will be agreed on with all researchers prior to the start of actual research. The data ownership agreement describes who owns the data, and whether and what rights will be transferred. Copyright and intellectual property rights will also be secured before any data is made public.
The data will be published for reuse under the Creative Commons license CC BY-SA 4.0. The project will be given the permission to share their data by research subjects, in an anonymized form.

3. Documentation and metadata

How will you document your data in order to make the data FAIR (findable, accessible, interoperable and reusable) for you and others? What kind of metadata standards, README files or other documentation will you use to help others to understand and use your data?

The research team members are committed to the FAIR data principles (Findable, Accessible, Interoperable, and Reusable), while at the same time following the principle "as open as possible, as closed as necessary."

To make our data **findable**, the project will archive the digital data in a publicly available repository, like IDA (https://ida.fairdata.fi/login) and use descriptive metadata as required and provided by

the data storage repository. Note that **metadata** can be made openly available in a Finnish or international data finder wherever possible, even when the dataset itself is not available).

The open research data and associated metadata will be assigned unique DOI (digital object identifiers) enabling (meta)data uniquely identifiable, and thus **accessible** and referenceable.

Research datasets are registered in Hanken's research database HARIS with the persistent identifiers for the (meta)data.

When possible, research partners will structure the metadata by using an appropriate, agreed-upon metadata standard format, and all variables be described by using the suitable, specific metadata standard and controlled vocabularies.

To make our data openly accessible, tools, software, and components used in the project will be available as open source as far as privacy allows.

To make the data accessible and **interoperable**, used formats will be based on open standards. When depositing data in the above repository, the project will ensure that the research data is migrated to new formats, platforms, and storage media as required by good open science practice to enable data sharing, reuse, and interoperability between researchers, institutions, organisations, and countries.

Files and folders will be versioned and structured by using a name convention consisting of project name, dataset name, and version information. Search keywords and subject headings from the KOKO Ontology will be provided to optimize data reuse possibilities.

For the data to be **reusable**, a creative commons license CC-BY 4.0 will be used for the project's outputs and open research data, free of charge for any user and without any embargo period, to ensure that they are shared with minimal restrictions, aside from attribution to the authors or creators.

Before submitting and depositing data in the chosen repository, all direct identifiers will be removed, indirect identifiers removed or categorized, and all sensitive personal information deleted to ensure that the data is properly and irreversibly anonymized. Any part of the data that cannot be anonymised will not be submitted for archival.

4. Storage and backup during the research project

4.1 Where will your data be stored, and how will they be backed up?

The research team members will follow Hanken's "Data security instructions for personnel."

During the active research period (when the research team members are still analysing the data), the data will be stored in and shared through such cloud-based information systems that are provided and maintained by Hanken (besides research team members' private computers), such as:

1. in individual researcher(s)' password-protected personal computers,
2. in password-protected joint-use computers in a room located physically at Hanken,
3. on memory sticks stored in locked closets/lockers of the researcher(s), and/or
4. in Hanken-provided network or cloud storages/drives (Hanken's network drives or Microsoft OneDrive for Business).

Hanken-provided systems do automatic backups. Therefore, the data are retrievable in case of human error or data corruption. In addition, manual backups of master data files will be taken regularly and always before any major file-format or data conversions.

The variables or data types which can directly identify a research participant, which are stored in a separate contact information file, and which represent sensitive personal information, will not be stored, at any stage, in other network or cloud-based storages than Hanken's own storages/network drives, or some of the storages indicated below:

1. individual research team members' password-protected personal computers,
2. password-protected computer or hard drive in a room located physically at Hanken, with restricted access, and
3. Hanken-provided network or cloud storages/drives/services.

4.2 Who will have access to the data during the research? Who will be responsible for access control? And how will secured access be controlled?

Right to access the data and data usage are controlled by the PI. The PI completes the list of users and all rights granted, and a procedure for withdrawing rights. Technical access control is provided by IT-services of Hanken. Data will be available to all research members of the project by using the "Specific people" option in Hanken's OneDrive portal, to keep control over who can be the authorized users.

Access control will be in line with the level of confidentiality involved. Data with direct identifiers, contact information, and sensitive personal data will not be sent between research team members by email – not even through Hanken's email system.

5. Opening, publishing and archiving the data after the research project

5.1 What part of the data can be made openly available or published? Where and when will the data and/or their metadata be made available?

In the spirit of open science, Hanken, as well as many research funders, recommends researchers to reuse existing archived datasets and, at the same time, to archive their research data in data depositories, so that other researchers can utilize and analyse the data later. A data archiving plan is part of research quality and transparency.

If you plan to archive research data, consider **what part of the data**, **where**, and **why** the data will/can be opened and published. Specify in the DMP that:

"Data will be available and cited in peer-reviewed international journals. The project ensures that all the publications will be immediately openly accessible. Copies of the publications will be uploaded and preserved in Hanken's institutional repository, DHanken.

After all direct identifiers are removed, indirect identifiers removed or categorized, and all sensitive personal information removed. The data will be archived until further notice for the purpose of allowing other researchers as well as scientific publication outlets to conduct further scientific analyses on the data. The part of data that cannot be anonymised will not be submitted for archival. The digital data and the associated metadata will be archived in a data storage repository, such as IDA (https://ida.fairdata.fi/login).

5.2 Will data with long-term value be preserved? If yes, where and for how long?

The research project do not consider the long-term preservation of the research data, i.e. a preservation time beyond 25 years.

6. Data management responsibilities and resources

Who will be responsible for specific tasks of data management during the research project life cycle? Also estimate the resources (e.g., financial, time and effort) needed for the data management tasks.

The project PI is responsible for the initial planning and execution of data management procedures. The researchers who have access to the data are responsible for the tasks of data collection,

data quality, data storage and backup, data documentation, data archival and and sharing. Hanken's research support unit will provide guidance for data management/stewardship and management of IPRs.

The research project team have allocated time to complete the data management tasks. Archival in the repositories is free of charge. The data management tasks altogether will take around 1 month during the 2-year project.