Plan Overview

A Data Management Plan created using DMPTuuli

Title: AI-Assisted Multicultural Interactions and Language Learning in Healthcare Work and Education (AIMIL)

Creator: Maria Frick

Principal Investigator: Maria Frick

Data Manager: Maria Frick

Affiliation: University of Oulu

Funder: The Research Council of Finland (former The Academy of Finland)

Template: Research Council of Finland

Project abstract:

The scientific and societal objectives are to 1) increase understanding of linguistically asymmetric multicultural interactions in healthcare, especially in patient work and workplace interactions of nurses, and their impact on patient safety and language learning in the Finnish healthcare sector; 2) identify and study effective AI-based solutions that facilitate nurses' multicultural interactions in the workplace and assist in language learning and practicing workplace interactions in educational environments before actual on-site training

ID: 29355

Start date: 01-01-2026

End date: 27-12-2028

Last modified: 09-11-2025

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

AI-Assisted Multicultural Interactions and Language Learning in Healthcare Work and Education (AIMIL)

1. General description of data

1.1 What kinds of data is your research based on? What data will be collected, produced or reused? What file formats will the data be in? Additionally, give a rough estimate of the size of the data produced/collected.

- Video-recorded face-to-face interactions at workplace and educational settings
- Video-recorded focus group and expert interviews
- Mutual Understanding Scale questions
- Word-associate tests
- Anonymous surveys

FILE FORMATS
Digital video data files generated will be in MPEG-4 (.mp4) format. MPEG-4 is an International Standards Organization (ISO) specification and this format is readable for most media players. Transcripts of the recordings are created using, e.g. Elan, and exported in .docx or .rtf files. The data from questionnaires is stored in .xlsx and/or .pdf format, and the Readme file as .docx. Managerial documents and project deliverables are stored in .docx, .xlsx and/or .pdf formats.

SIZE ESTIMATE = 2 Tb

1.2 How will the consistency and quality of data be controlled?

Transcripts are checked by at least 2 people before release.
Files are named in a consistent manner.

2. Ethical and legal compliance

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

ETHICS AND LEGAL COMPLIENCE.
Participants receive information sheets with details of the project and their rights according to GDPR and Finnish legislation. They sign a contract agreeing that the recordings will be made available for research during and after the project. For children under 15, parental signature is required as well. Oral information about the study will be given to participants of all ages, and they will have the opportunity to ask questions. Information sheets and consent forms will be made available in languages accessible to all participants. Participants have the right to withdraw from the study without

negative consequences.

GDPR

The collection and storage of personal data is necessary, because the objective of the study is to study natural, unaltered conversations and embodiment, such as facial expressions, which requires unaltered video footage.

Personal data / sensitive data is processed on the following basis, which is based on Article 6(1) of the General Data Protection Regulation:

☒ performance of a task carried out in the public interest or in the exercise of official authority vested in the controller:

☒ scientific or historical research purposes or statistical purposes

The following sensitive personal data may occur in the video recordings, but it is not the intended object of study:

☒ Racial or ethnic origin

☒ Political opinions

☒ Religious or philosophical beliefs

☒ Trade union membership

☐ Genetic data

☐ Biometric data for the purpose of uniquely identifying a natural person

☒ Health

☒ A natural person’s sex life or sexual orientation

☒ Data regarding criminal convictions and offences or related security measures

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

The intellectual property of the data generated will remain with researchers, but will be also handed to the University of Oulu via a signed contract.

Data from the University of Oulu Kikosa Collection can be shared with selected researchers in EU. The selection of researchers is based on a research plan, which should be sent to maria.frick@oulu.fi. Researchers using the data are bound by the regulations stated in the original collection.

3. Documentation and metadata

3. How will you document your data in order to make the data findable, accessible, interoperable and re-usable 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?

DATA DOCUMENTATION AND QUALITY

Metadata /transcription files are created of each recording, stating the length and format of the recording, and information about the speech situation and the speakers (e.g. age, gender, language skills). Recordings are given a name that will help identify them. Participants are given pseudonyms (codes) that will link background information to the speakers in the transcripts.

README-files and other instructions will be written if necessary for better consistency and access to the data.

4. Storage and backup during the research project

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

The data, (as well as managerial documents and project deliverables) is added to the University of Oulu Kikosa collection, which is currently stored on a limited-access folder on the university server. Sensitive data is password protected. The data may be stored at a more secure location later, if such an opportunity arises.

4.2 Who will be responsible for controlling access to your data, and how will secured access be controlled?

Access is controlled by the PI, Maria Frick.

Data can only be made public in accord with the contract signed by participants.

During the research project, data will be handled only by members of the project team and related persons (e.g. students working under supervision of the team members).

When accessing the data from outside the EU, project members and collaborators will use a University of Oulu VPN-connection or EU-based, password-protected platforms such as the CSC Sensitive Data (SD) Desktop.

Parts of the data (videoclips, pictures etc.) may also be presented publicly and published on research related occasions such as scientific conferences, journals or teaching, and parts of the data can shown to project collaborators within and outside the EU, using a University of Oulu VPN-connection or a password protected EU-based platform such as Zoom, Panopto or CSC’s SD Desktop.

Participants’ full names will not be published.

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, or its metadata, be made available?

Detailed metadata of the Kikosa collection is stored in an Excel-file, and also published in http://urn.fi/urn:nbn:fi:csc-kata20190315143424155015

The data cannot be made openly accessible because it contains personal data and may contain sensitive data. The intended purpose of the data collection remains in scientific research, which is why it's use is restricted.

5.2 Where will data with long-term value be archived, and for how long?

The data will be added to the University of Oulu Kikosa collection for later use and permanent storage.

6. Data management responsibilities and resources

6.1 Who (for example role, position, and institution) will be responsible for data management?

University lecturer Maria Frick, University of Oulu
Coordinator, Researchers, Research assistant

6.2 What resources will be required for your data management procedures to ensure that the data can be opened and preserved according to FAIR principles (Findable, Accessible, Interoperable, Re-usable)?

Staff time has been allocated in the proposed budget to cover the costs of preparing data and documentation for archiving. The University of Oulu offers resources (storage space and technical support) for the management and permanent deposition of the data.