Plan Overview

A Data Management Plan created using DMPTuuli

Title: PhD- Images of the future of young people

Creator: Pasi Keski-Pukkila

Affiliation: University of Turku

Template: General data management plan - University of Turku

Project abstract:

This PhD is examining the vocational images of the future of young people. The aim is to understand what kind of fears, hopes and expectations are visible in these images and how these images have initially formed. The goal of the research is to reveal assumptions, beliefs and power structures that are affecting vocational decision-making process and the future-oriented thinking of young people.

ID: 27977

Start date: 01-01-2024

End date: 31-12-2027

Last modified: 30-04-2025

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

PhD- Images of the future of young people

1. General description of the 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 and collected.

1. General description of the 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.

This research is based on **newly collected qualitative data**, gathered through **workshops** and semi structured **in-depth interviews**.

- **Workshops (16–18-year-olds)**: Participants will produce personal writings, group work materials (posters), and recorded discussions.
- **In-depth interviews (18+ year-olds)**: 25 semi-structured interviews, each lasting approximately 90 minutes, will be recorded and transcribed.

The data will be stored in the following **file formats**: Excel, PDF, DOCX, JPEG, and MP3 and WAV.
The estimated size of the collected data is:

- **Personal writings**: ~50 pages
- **Workshop recordings**: ~17 hours
- **Interview recordings**: 25 x 90 minutes (approx. 37.5 hours)

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

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

To ensure the consistency and quality of data, the research follows standardized data collection and processing methods:

- **Workshops**: Participants will respond to a structured set of prompts for personal writings, ensuring all responses address similar themes. Group work will follow a predefined facilitation process with clear instructions and standardized materials (posters). Recorded discussions will be guided by a fixed set of discussion themes.
- **In-Depth Interviews**: A semi-structured interview guide will be used to ensure all interviews cover key themes while allowing flexibility for follow-ups. All interviews will be recorded in a consistent format and quality setting to facilitate accurate transcription.
- **Data Processing & Documentation**: Workshop and interview recordings will be transcribed **verbatim** to maintain accuracy. A systematic coding scheme will be used to analyze data consistently, and metadata (participant ID) will be recorded to track data sources reliably.

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.)

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

This research involves the collection of **personal data**, including participants' names, ages, schools, and gender. To ensure compliance with the **General Data Protection Regulation (GDPR)** and national data protection laws, the following measures will be taken:

- **Anonymization and Pseudonymization**: All personally identifiable information will be pseudonymized during the data processing phase and, where possible, fully anonymized before data sharing or publication.
- **Informed Consent**: All participants (both workshop and in-depth interview participants) will provide informed consent before data collection.
- **Guardian Notification**: The guardians of participants under 18 will be **informed** about the study, ensuring transparency.

- **Secure Storage & Access Control**: Personal data will be stored securely, with restricted access, in compliance with university and GDPR guidelines. Supervisors will not have access to any raw personal data or identifiers at any stage of the research.
- **Institutional Compliance**: The research follows the **data protection policies and ethical guidelines of the university of Turku** ensuring legal and ethical best practices.

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

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

This research involves the collection of personal data, including participants' names, ages, schools, and gender (optional). To ensure compliance with the General Data Protection Regulation (GDPR) and national data protection laws, the following measures will be taken:

- **Anonymization and Pseudonymization:**
  All personally identifiable information will be pseudonymized during the data processing phase and, where possible, fully anonymized prior to data sharing or publication. Pseudonymization reduces identifiability by replacing direct identifiers with codes, while anonymization irreversibly removes identifying information. These measures ensure that individuals cannot be directly or indirectly identified from the data.
- **Informed Consent:**
  All participants (both workshop and in-depth interview participants) will provide informed consent before data collection. This consent process will ensure participants are fully aware of the study's objectives, the data collection methods, and how their data will be used. The consent form, data protection statement and notification of the research to the guardians of minors will be sent at least one week before the workshop. In the case of individual interviews, participants will be provided with a consent form to review in advance whenever possible, ideally at least a week prior to the interview. However, if a participant is recruited unexpectedly or on short notice, it may not be feasible to provide the consent form a week in advance. In such cases, efforts will be made to ensure that the participant fully understands the nature of the study and their involvement before proceeding with the interview. Clear, verbal explanations will be given to guarantee informed consent is obtained, even in situations where written consent cannot be reviewed in advance.
- **Secure Storage & Access Control:**
  Personal data will be securely stored in compliance with GDPR and institutional guidelines. Access will be restricted to authorized individuals only, ensuring that data is handled responsibly and kept confidential.
- **Institutional Compliance:**
  This research will adhere to the data protection policies and ethical guidelines of University of Turku, ensuring that all data handling and sharing practices meet the legal and ethical standards required for the study.

3. Documentation and metadata

How will you document your data in order to make it 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?

3.1 How will you document your data in order to make it 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?

This research involves the collection and organization of data, including the written letters by participants, recorded discussions and posters compiled during the workshop. To ensure the data is findable, accessible, interoperable, and reusable, the following steps will be taken:

- **Data Organization:**
  Data will be systematically organized as it is collected. For example, participant letters will be categorized by topic, and recorded discussions will be grouped by school, theme and group. This will ensure that the data is easy to navigate and retrieve when needed.
- **Metadata:**
  Basic metadata will be included for each dataset, such as titles, collection dates, and relevant themes or categories. This will help make the data easily searchable and identifiable.
- **File Formats:**
  The data will be stored in standard formats, such as text files for participant letters and audio files (e.g., MP3, WAV) for recorded discussions. These formats are widely accessible and allow easy use and analysis.

4. Storage and backup during the research project

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

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

During the research project, the data will be stored primarily on my personal computer, which will be password-protected to ensure its security. Additionally, the data will be uploaded to **Seafile**, a university-provided secure storage system, where it will be regularly backed up through an automatic backup system. This ensures that the data is protected against loss or corruption. Wherever the data is stored, it will always be password-protected to prevent unauthorized access. In the case of sensitive data, additional precautions will be taken to safeguard it, in line with institutional security protocols.

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

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

I will be the sole person responsible for controlling access to the data. Access will be restricted exclusively to myself, and all data will be stored in a password-protected environment. No one else will have access to the data unless I explicitly grant permission. This ensures that the data remains secure and protected throughout the research process.

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?

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?

A selected set of data, specifically from participants who have given consent, will be made publicly available after the completion of the research. This data will be archived in the Social science Data Archive at Tampere University for long-term preservation. The data will be anonymized and pseudonymized to ensure privacy protection before being released publicly.

Once the data is archived in the repository, it will be made findable through the archive’s searchable platform. Although the exact process of locating the data may depend on the archive’s specific structure, it is anticipated that the data will be assigned a DOI (Digital Object Identifier) and tagged with relevant metadata, including titles, keywords, and themes, to improve discoverability. This will allow researchers and other users to easily locate the dataset using search terms related to the research topics.

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

See 5.1

6. Data management responsibilities and resources

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

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

I, **Pasi Keski-Pukkila**, as a PhD researcher at the **University of Turku**, will be primarily responsible for managing the data throughout the research project. This includes the collection, storage, backup, and eventual archiving of the data. I will also ensure that the data management practices adhere to ethical and legal guidelines.

My supervisors, **Sanna Ahvenharju** and **Katriina Siivonen**, will provide guidance and oversight as needed but will not directly manage the data.

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)?

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)?

To ensure that the data can be opened and preserved according to the **FAIR principles** (Findable, Accessible, Interoperable, Re-usable), the following resources will be required:

- **Data Storage and Backup:**
  Secure storage of the data will be managed through **Seafile**, the university-provided system, which offers encrypted storage and an automatic backup system. This ensures the data remains accessible and protected during the research process.
- **Metadata and Documentation:**
  I will add relevant metadata to the data to ensure it is Findable and Reusable. While I have not yet decided on the specific software or tool to use for metadata creation, I will follow established standards such as **Dublin Core** and **DataCite**. Additionally, I will create a **README file** to document the data structure, collection process, and any relevant details to support Reusability.
- **Data Formats:**
  The data will be stored in widely used formats, such as **CSV** for quantitative data and **text or audio files** for qualitative data. This ensures that the data is Interoperable with common analysis software and can be reused by other researchers.
- **Archiving and Publishing Platforms:**
  After the research is completed, the data will be archived in the **Social science Data Archive** at **Tampere University** for long-term preservation. The archive will assign a **DOI** (Digital Object Identifier) to the dataset, and the data will be tagged with relevant metadata to improve discoverability.
- **Institutional Support:**
  Support from **University of Turku** IT services and data management professionals will be available to assist in ensuring that the data is properly archived and meets the required standards for public access.