Plan Overview

A Data Management Plan created using DMPTuuli

Title: Strengthening collaboration towards effective and integrated services for children, young people and families with children in Southwest Finland and West Uusimaa

Creator: Katri Aaltonen

Principal Investigator: Silja Kosola, Katri Aaltonen, Johanna Kallio

Contributor: Margus Pruel, Hanna Rättö

Affiliation: University of Turku

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

Template: Research Council of Finland

ORCID iD: 0000-0002-1737-264X

Project abstract:
CoWest a) strengthens register-based research of integrated social, healthcare and education services, b) produces evidence on how service systems promote the health and wellbeing of children, youth and families with children, c) promotes transdisciplinary research-practice co-creation for knowledge management and practical work in Southwest Finland and West Uusimaa. The project provides a framework for register-based research by integrating data mined from regional information systems with national datasets. It addresses the needs of regional knowledge management by supporting the development of indicators related to the organization and effectiveness of services. The regional database created enables research beyond the scope of the project itself and even before the accumulation of national data. Future research projects can be connected to the regional database, making it a novel register dataset from an international perspective.

ID: 29961

Start date: 01-01-2026

End date: 31-12-2028

Last modified: 23-01-2026

Grant number / URL: 373060

Strengthening collaboration towards effective and integrated services for children, young people and families with children in Southwest Finland and West Uusimaa

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.

The project reuses individual-level register data primarily collected to document client and patient interactions in routine service and clinical settings. This includes nationally collected register data held by Statistics Finland (SF), Finnish Institute for Health and Welfare (THL) and Social Insurance Institution (Kela), as well as regional / local databases of wellbeing services counties (WSC) and municipalities encompassing data on education, social and healthcare services. Furthermore, interview data are collected during the project, including interview recordings and informed consent forms collected from participants, verbatim transcripts of recordings, administrative documents, and participant lists for interview administration. Data, file formats and sensitivity are described in the Table below.

Data type	File format, size	Personal data, sensitivity	Controller	Storage and handling during the project
A. National register data (Statistics Finland, Kela, THL)	.csv, .dta, .sas7bdat >100GB	Pseudonymised personal data, sensitive (health data)	Varha, UTU, Luvn	Statistics Finland Fiona
B. Client information systems data at Wellbeing Service Counties (WSC)	TBD >100GB	Personal data, sensitive (health data)	Wellbeing Service Counties	Wellbeing Service Counties' IT systems
C. Educational services information systems at the municipalities	TBD >100GB	Personal data, sensitive	Municipalities	Municipalities' IT systems
Register data structured from B +/- C	.csv, .dta, .sas7bcat >100GB	Pseudonymised personal data, sensitive (health data)	Each Wellbeing service county controls its own data	HUS Acamedic, Varha Atolli
Linked register data combining A + B + C	.csv, .dta, .sas7bcat >100GB	Pseudonymised personal data, sensitive (health data)	Varha, UTU, Luvn	Statistics Finland Fiona
Interview recordings	.avi, .mp4 >10GB	Personal data, non-sensitive	Varha, UTU	Varha Atolli or Varha's servers
Interview administrative documents	.pdf, .xlxs, .docx <100MB	Personal data, non-sensitive	Varha, UTU	Varha Atolli or Varha's servers
Interview transcripts	.pdf, .txt, .xlxs, .docx <100MB	Anonymised data	Varha, UTU	Varha Atolli or Varha's servers, UTU servers

Table 1. Overview of data used in the project

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

Formal version control is applied for all types of data. Original/raw data are kept separately to ensure that the accuracy of data is maintained and no data are accidentally changed. All users of data will be trained to implement quality control. All quantitative data handling and analyses will be conducted using log files. Interview transcriptions are checked by another person than the transcriber.

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.)

Data Protection is based on legal obligations as described in the EU General Data Protection Regulation (GDPR) and national Law, and as set forth in the Data Protection policies of participant and host institutions. Ethical review statements will be sought for each subproject initiated during the project duration from relevant Ethics Committees. Permissions to use secondary register data for research purposes are applied for from the relevant regional and national authorities, and used in compliance to their instructions and policies, with Data Protection Impact Assessment (DPIA) as part of the process. The data includes identifiable and sensitive personal data. The processing basis for register data is scientific research, in which case informed consent is not required according to Finnish legislation. WSCs are allowed to use their own data for knowledge management purposes based on national legislation (L 552/2019 §41). Municipal data of educational services, and permissions to link data with other registers, are applied for from the municipalities. Data are pseudonymised by data holders or upon receipt and all data handling and analyses are conducted with pseudonymised data. The processing basis of interview data is subject's freely given, specific, informed and unambiguous consent. Minimal data on participants is collected, and all results are reported in a manner that does not allow identification of individuals. Lists of participants for interview

administration include identifiable non-sensitive personal information and are kept separately from other data, and deleted after data collection and verification processes. Privacy Notice of Processing Personal Data is given to all interview participants. Interview recordings will be deleted once transcripts have been produced and checked for consistency.

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

Rights to use regional/local data produced and collected in preceding research projects LAREK, Pikkuiset and Sotepala remain separate, as originally agreed between the data holders and parties conducting each research project. A new, jointly controlled register-based dataset (Varha, UTU and Luvn) will be created during the project. The rights to use the jointly controlled data are managed through the permission processes with Statistics Finland, Findata, as well as the regional and municipal data holders. A written agreement between Varha, UTU and Luvn will be made to agree on the rights and responsibilities between the data controllers in relation to the joint register data. Possibilities of producing ready-made datasets from the municipal and WSC data for further research use via Statistics Finland will be explored with the municipalities, WSCs and Statistics Finland. Interview data will be controlled by UTU or Varha, and if needed, through an agreement on joint controllership. Open access is granted to reports, replication codes, metadata, scientific publications and other publications using as broad user licenses as possible, typically CC BY 4.0.

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?

Versioning is used for all files, supported by, e.g., GitHub, and continuously updated README-files in storage root and each sub-folder, with descriptions of naming conventions, purpose and files in that storage organisation level. The institutions that manage secondary register data provide the general description of data (e.g., metadata), questionnaire, codebook, and user guides on their websites. Metadata for ready-made datasets is developed in collaboration with Statistics Finland.

4. Storage and backup during the research project

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

Wellbeing services counties and municipalities data are stored within their institutional IT systems. Data used for research purposes are stored and used in operating environments under the Act on the Secondary Use of Health and Social Data registered in the Astori public register maintained by The National Supervisory Authority for Welfare and Health (Valvira). The operating environments used in this project are Statistics Finland's Fiona, Helsinki University Hospital (HUS) Acamedic and the Wellbeing Services County of Southwest Finland Atolli. Interview data are stored on servers provided and technically administered by Varha, including backups, access control limited to project researchers, and security.

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

Right to access the data in Fiona is controlled by the PIs and the appointed data managers, and technical access control is provided by Statistics Finland. Right to access the data in Atolli is controlled by PI Johanna Kallio and the appointed data managers, and technical access control is provided by Auria Clinical Informatics. Right to access the data in Acamedic is controlled by the PI Silja Kosola and the appointed data managers, and technical access control is provided by Helsinki University Hospital. All project data have access control limited to project researchers with appropriate permissions to use data. Access to data will be documented. The PIs will be at any point able to tell who has access to what data.

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?

Possibilities of producing ready-made datasets from the municipal and WSC data for further research use via Statistics Finland will be explored with the WSCs, municipalities and Statistics Finland. Annotated syntax can be made openly available upon publication of each scientific research article. Possibility to share interview data through for example Finnish Social Science Data Archive will be dependent on whether data can be anonymised and conditional to explicit consent for archiving by the interviewees.

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

Wellbeing services counties and municipalities archive their own data according to their standard operating processes. Research data used in the operating environments Fiona, Atolli and Acamedic are destroyed according to the agreements made with data authorities and data holders. Ready-made datasets will be managed by Statistics Finland as per agreements between Statistics Finland and the data holders.

6. Data management responsibilities and resources

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

The overall data management responsibility, including DMP implementation and annual review during the project life cycle lies on the PIs (Johanna Kallio, Silja Kosola, Katri Aaltonen). PIs are responsible for ensuring that personnel working with project data have applied appropriate permissions, and are committed to using data as per legal obligations, ethical guidelines, and institutional data protection policies of participant and host institutions. The PI's will delegate the day-to-day data management responsibilities to data managers working in the project.

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)?

The budget for data acquisition, management and remote use, as well as interview administration and transcription are included in the research plan. Data management is part of the tasks of the PIs and personnel working on the project, the salaries of whom are included in the research plan. The project benefits from the support of FIRE infrastructure and INVEST at UTU, and project collaborators Statistics Finland and the Finnish Institute for Health and Welfare.