

# WP 3.3 QUALITATIVE SURVEY DATA

## Contributors

- Krista Lagus, UH-CSDS
- Joni Oksanen, UH-CSDS
- Maria Valaste, UH-CSDS
- Tuukka Oikarinen, UH-CSDS
- Maria Litova, UH-CSDS
- Katja Laine, UH-CSDS
- Ida Toivanen, JYU
- Tomi Oinas, JYU
- Sakari Taipale, JYU
- Jari Ojala, JYU
- Tanja Välisalo, JYU
- Jani-Matti Tirkkonen, UEF
- Jaakko Peltonen, TAU

## ANALYSIS PIPELINE

The lack of easy-to-use analysis tools for Survey open text questions poses a challenge for researchers. Survey answers have rich structural and statistical contexts that sometimes extend to population representativeness. The objective here is to design a shared, easy-to-use and statistically sound pipeline for analysing the qualitative text data in surveys and for linking it with central aspects and concepts of the survey data set.

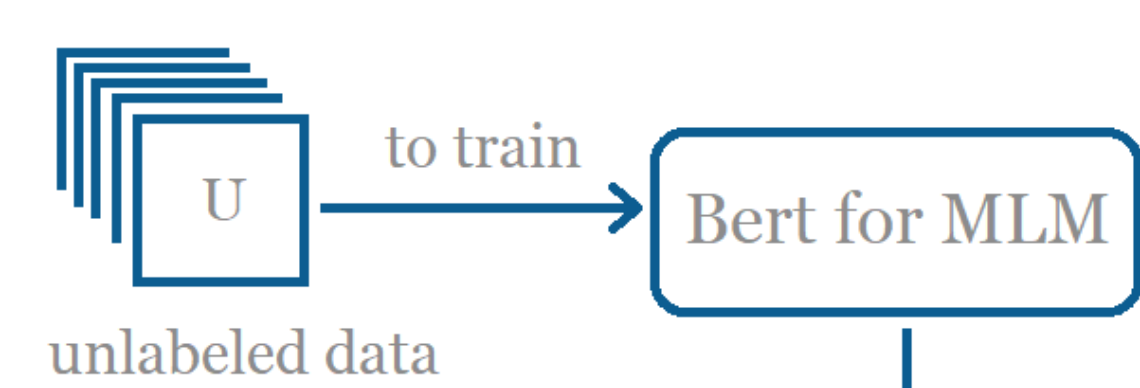
## PILOT STUDIES

We work simultaneously with several pilot surveys: Tech in eldercare, Loneliness, Child barometers, Faculty social wellbeing, and Everyday life in Varkaus. Two data sets were obtained from the Finnish Social Science Data Archive (FSD), two from research projects in progress, and one is yet to be obtained. Data set sizes vary from below 300 to over 4000 text answers. Various analysis approaches will be explored on these data sets, in order to design the pipeline. We also examine interfaces and data permission processes.

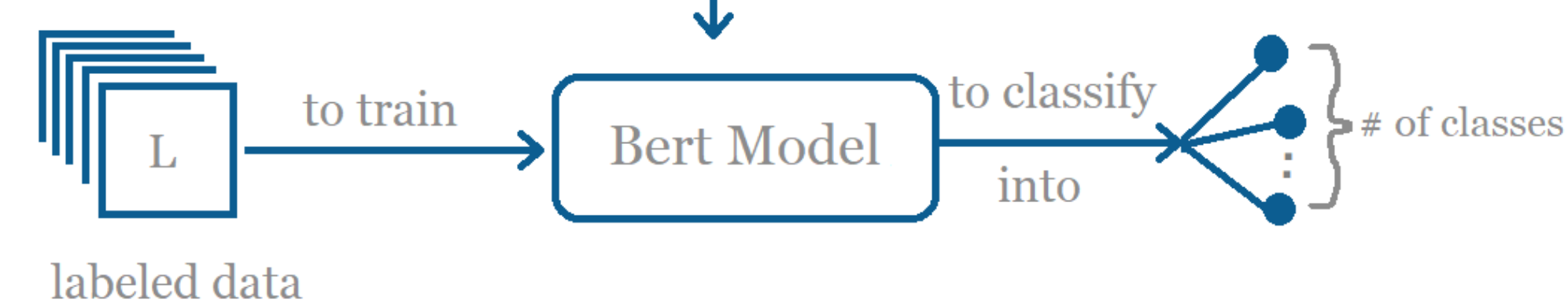
## TECH IN ELDERCARE

*JYU and everyone.* Pilot case. The data were collected in 2019 and 2021 from two studies conducted by Centre of Excellence in Research on Ageing and Care (2018–2025). The objective was to obtain information about the workers employed in elder care, their working environment and conditions, technology used in care work, such as devices and services, and the effects of digitalization on work practices. The questionnaires had 62 questions, including two open-ended ones. For one open-ended question ("What kind of emotions related to the use of technology have been present in your work during the last week?") around 4000 answers were obtained and used to train a FinBERT model for sentiment analysis. The data was also used for semi-supervised deep learning for text classification (Toivanen, in press).

### Pretraining



### Finetuning



## LONELINESS

*UH-CSDS.* This survey conducted by a Finnish newspaper Helsingin Sanomat (HS) in 2014, charted Finnish experiences of loneliness and surveyed the circumstances of loneliness and overall satisfaction with life. Dataset is openly accessible from the FSD. The survey included one open-ended question "What does loneliness feel like?". The open-ended responses to this question are only available in Finnish. Other themes of the survey included The dataset has 27,851 responses in total and 3491 open-ended responses (12.5%).

Survey respondents reflect the readership of HS, most of the respondents are working aged (20-40 years old) adults from Southern Finland. On average, the respondents self-assess themselves as middle class, supported by their gross income and ability to economically manage their expenses. There were roughly 20% more women than men in the respondents.



## FACULTY SOCIAL WELLBEING

*UH-CSDS.* In Spring 2022, The Faculty of Social Sciences at University of Helsinki pilots the use of crowdsourcing to improve the social wellbeing and equality at the Faculty. Both survey and discussion forum data are collected anonymously in the open-access *otakantaa.fi* platform. The crowdsourcing is advertised in Faculty internal email lists, in University premises and in social media. The dataset (as of 23.5.2022) consists of 211 answers to survey open questions and 15 text messages from the discussion forum that comment on the proposed Code of Conduct. The survey contains questions on background information, 13 closed questions related to subjective experiences of equality or social wellbeing at the faculty and 11 open-ended text questions.



UNIVERSITY OF HELSINKI  
FACULTY OF SOCIAL SCIENCES



UNIVERSITY OF JYVÄSKYLÄ



UNIVERSITY OF  
EASTERN FINLAND



Tampere  
University

## CHILD BAROMETERS

*UH-CSDS.* Child Barometers study the everyday lives, eg., levels of trust, well-being and exercise of 6-year-old Finnish children. The samples for different study years have been randomly selected from the population system of the Digital and Population Data Services Agency and the target population was children aged 6 living in mainland Finland. For study years 2016 and 2020 sampling method was simple random sampling and for study year 2018 multistage sampling. For years 2016 and 2020 weights are included.



Surveys included background variables and open-ended questions. The surveys were conducted by the Office of Ombudsman for Children and the datasets are archived at the FSD with open access.

## EVERYDAY LIFE IN VARKAUS

*UEF.* The survey was conducted in 2016 by researchers at the University of Eastern Finland. The objective was to study a typical town in Eastern Finland that has gone through structural change. The respondents were randomly selected people from Varkaus between ages 18 and 85 years, yielding a total of 1970 answers. The survey asks about: background information, work, family and other relationships, wellbeing, free time, political views, and the use of services. There are three open questions, where the participants are asked to describe their family and what is most important in their life.