EUROPEAN RESEARCH COUNCIL (ERC): DATA MANAGEMENT PLAN (DMP)

CRISP

Project Number: 803239
Project Running Time: 01 January 2019 – 31 December 2023 (60 months)

Version 1
Date: 11 June 2019 (6-month deliverable D1.1)
Scheduled for updates:
30 June 2021 (30-month first scientific report)
31 December 2023 (60-month final scientific report)

PI: Assoc. Prof. Anja Leist (anja.leist@uni.lu), University of Luxembourg

PROJECT SUMMARY

Cognitive impairment and dementia have dramatic individual and social consequences, and create high economic costs for societies. In order to delay cognitive aging of future generations as long as possible, we need evidence about which contextual factors are most supportive for individuals to reach highest cognitive levels relative to their potential. At the same time, for current older generations, we need scalable methods to exactly identify individuals at risk of cognitive impairment. The project intends to apply recent methodological and statistical advancements to reach two objectives. Firstly, contextual influences on cognitive aging will be comparatively assessed, with a focus on inequalities related to educational opportunities and gender inequalities. This will be done using longitudinal, population-representative, harmonized cross-national aging surveys, merged with contextual information. Secondly, the project will quantify the ability of singular and clustered individual characteristics, such as indicators of cognitive reserve and behaviour change, to predict cognitive aging and diagnosis of dementia. Project methodology will rely partly on parametric 'traditional' multilevel- or fixed-effects modelling, partly on non-parametric statistical learning approaches, to address objectives both hypothesis- and data-driven. Applying statistical learning techniques in the field of cognitive reserve will open new research avenues for efficient handling of large amounts of data, among which most prominently the accurate prediction of health and disease outcomes. Quantifying the role of contextual inequalities related to education and gender will guide policymaking in and beyond the project. Assessing risk profiles of individuals in relation to cognitive aging will support efficient and scalable risk screening of individuals. Identifying the value of behaviour change to delay cognitive impairment will guide treatment plans for individuals affected by dementia.

DATASET SUMMARY

The project relies on secondary analysis of collected data. All datasets intended for use are available to the scientific community for free. The survey websites that are referred to below specify the conditions of application for and use of the data. The project firstly uses aging surveys harmonized with the Health and Retirement Study (HRS) across the world. Currently foreseen is the analysis of

- The Health and Retirement Study of the U.S.
- The Survey of Health, Ageing and Retirement in Europe: list of available DOIs
- Korean Longitudinal Study of Aging
- English Longitudinal Study of Ageing: DOI associated with bibliographic citation
- The SAGE survey by the WHO

Over the project running time, further harmonized surveys will be added, for a full list of available countries, their coverage, cohorts and years of study, see The Gateway to Global Aging initiative.

Additionally, country-specific datasets with a focus on aging will be used of surveys that are part of the IALSA network. Currently foreseen is the analysis of

- Longitudinal Aging Study Amsterdam
- The Health, Aging and Retirement Transitions in Sweden (HEARTS) study

All of the used datasets already adhere to the FAIR principles:

1. Findable: the survey websites give information how to access the data. Many datasets have DOIs; all available DOIs are listed here.

2. Openly accessible: Details on how to apply for and use the data have been mentioned above.

3. Interoperable: The HRS surveys strive for harmonization across studies which guarantees interoperability; efforts in the scientific community to make surveys of IALSA studies interoperable will be adhered to.

4. Increase data reuse: All surveys intended for use have different measures to increase data re-use; among those lists of publications that used the data; how-to instructional working papers and guidelines to use the data; provision of survey details as cohort profiles e.g. in the International Journal of Epidemiology, and to harmonization efforts as done in the Gateway to Global Aging.

The data from aging surveys will be merged with macro-level indicators for countries and cohorts, that are provided by different institutions and projects free of charge and with minimal requirements (e.g. not even user name and professional affiliation are needed for the download of data). Examples of those data are: World Bank Open Data, World Income Inequality Database, Barro-Lee Educational Attainment Dataset.

As for allocation of resources and data security that the Open Research Data Pilot requires, the project relies on already stored and publicly accessible data. There is no need for new storage/making the survey data accessible.

RESEARCH DATA SUMMARY

All publications will be made open access in accordance with the ERC's guidelines. On top of that, the Open Research Data Pilot's guidelines suggest that all research data be made accessible. For the CRISP project, this is specifically important to programming code and other associated output of the project for replication of the findings etc. Project publications will be supplemented by (Stata, R) programming code where possible to make replications easier. Exceptions will be made only in well justified cases where ethical considerations need to be prioritized against open data (detailed considerations can be found in the document *Reply to the ethics review* of 11 October 2018). All programming code will be submitted where possible directly with manuscripts as supplementary files, and will be stored where relevant directly on the project's website or in the host institution's repository ORBilu.