DATA MANAGEMENT PLAN 03/2023

Unity is strength: towards an integrative, dynamic, and resource-based perspective on violence, sexual harassment, and bullying as forms of interpersonal mistreatment at work.

Creator: Charlotte Franckx
Promotor: Elfi Baillien
Affiliation: KU Leuven (KUL)
Template: KU Leuven BOF-IOF

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

During the course of the project, we will: (1) make use of existing data, more specifically of published research, existing research models and theories, and (2) collect, analyze, and document new data. The project entails (WP1) qualitative data (manuscripts in MS Office), and (WP2) cross-sectional and three-wave quantitative data (survey data in .csv) from employees. The datasets will be saved as a datafile, both in CSV and in IBM SPSS. In addition, we will save the syntax (.sps file) and document the metadata.

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*	*ONLY FOR PHYSICAL DATA*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
WP1 Qualitative study: interviews with 40 subject matter experts	Qualitative data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☒ Sound ☐ Numerical ☐ Software ☐ Model ☒ Textual ☐ Other	Audio-recorded and stored in uncompressed WAV format; transcribed into MS Word format	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
WP1 Analyses	Results from qualitative analyses	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Model ☒ Textual ☒ Software ☐ Other	.docx, .odt or .txt	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
WP2 Longitudinal quantitative study via survey; +/- 2500 workers	Quantitative data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Model ☐ Textual ☒ Software ☐ Other	Database (.sav)	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
WP2 Validation of new resources scale via data collection service; 150 employees	Quantitative data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Model ☐ Textual ☒ Software ☐ Other	Database (.sav)	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
WP2 Analyses	Results from statistical analyses	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Model ☐ Textual ☒ Software ☐ Other	Database (.sav)	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Personal data	(name, e-mail, profession…)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Model ☒ Textual ☐ Software ☒ Other	Database (.sav) Excel files (.xlsx)	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:
Reused data will mainly constitute existing research and theories. These will be sourced from all kinds of journal/ interdisciplinary databases and will be identified via citations and reference lists in all the research papers and documents.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.
Yes, human subject data (Provide SMEC or EC approval number below). We will collect qualitative (personal) data from subject matter experts and (personal) survey data from employees. We got ethical approval for the project from SMEC (KULeuven) on 31st of August 2022 – code G-2022-5334-R2(MIN).

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).
Yes (Provide PRET G-number or EC S-number below): At the recruitment phase of study 1 (semi-structured interviews), interested participants will be able to contact us by e-mail after seeing the advertisement for the study through the research team's social networks and in specialized magazines. As a result, we will be informed of their e-mail addresses and names. These will only be used for (1) contacting the participants to provide them with information on the study and an informed consent file prior to the interview and (2) to contact participants who wish to receive a summary of the study's results. Names will not be stored; in the transcriptions each respondent will receive a number code in the transcription file. The recordings of the interviews will be stored in a secured file, separated from the transcription. They will be deleted after transcriptions/coding of the data has been performed. Also e-mail addresses will be stored in a secured file, separately from the subsequent interview transcriptions. This file will be permanently deleted after (2) has been met. Study 2 will be conducted through a data collection service handling the ethical procedure regarding personal data themselves: these will not be shared with the researchers. For study 3 (electronic surveys), the information letter and informed consent will immediately be displayed. We will also ask participants for their e-mail addresses at the registration stage of the survey. This e-mail addresses will be used for (1) linking answers of the three measurement points, together with a random code, (2) contacting participants who won a voucher if they registered for this, and finally (3) for contacting participants who wish to be informed about the results of the study. Regarding these last two purposes, participants will be asked whether they are interested in participating in the voucher lottery and in receiving the study results. They will have the possibility to fill in their e-mail address if interested. These will be the only three purposes for which e-mail addresses will be used. To ensure that they remain protected, the e-mail addresses will be stored in a separate and secured file on the researcher's personal and secured KU Leuven OneDrive for Business account. This file will be permanently deleted after distribution of the vouchers (2) and study report (3). Also the final data file with linked responses by participant will not refer to the e-mail addresses. Finally, also some socio-demographic data will be collected throughout the three studies. This includes age, gender, job information (function, sector of employment, whether or not the participant has a managing position). N.B.: audio-records and transcriptions of the interviews will be made for study 1. However, the questions of the interviews are formulated in such a way in the interview protocol that we impede possibilities to trace back the identity of the participant.
Approved by SMEC, G-2022-5334-R2(MIN).

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or da ta type where appropriate.
No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.
No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.
No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable , for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information Is recorded).
For every work package, metadata will be generated. First and foremost, a README text- (or Word-) file will be created according to the KU Leuven template. Consequently, the following information will be provided:

GENERAL INFORMATION
- G01. Names of file(s) or dataset(s) that this README file describes
- G02. Date of creation/last update of the README file
- G03. Name and contact information of Principal Investigator
- G04. ORCID of Principal Investigator
- G05. Institution of Principal Investigator
- G06. Contact of other person at KU Leuven that has access to the dataset

G07. Description of the dataset
G08. Keywords (author defined)
G09. Thesaurus or controlled vocabulary keywords
G10. Thesaurus or controlled vocabulary used in this README
G11. Language(s) used in the dataset
G12. Other involved researchers

PROJECT INFORMATION
P01. Project information
P02. Project abstract
P03. Project funder: Name of funder, type of grant, grant number

FILE OVERVIEW
F01. Number of files described by the README-file
F02. List with names of files, description, date of creation of file
F03. File formats
F04. Software used to generate the data
F05. Software necessary to open the file
F06. Relationship between the files
F07. Which version of the dataset is this? Date of this version?
F08. Information about the dataset versions and reason for updates
F09. Naming conventions for file names

STORAGE INFORMATION
S01. Where are the data stored?
S02. Links to other available locations of the dataset (e.g. repository)

METHODOLOGICAL INFORMATION
M01. Date (beginning-end) and place of data collection
M02. Aim for which the data were collected
M03. Data collecting method
M04. Information about data processing methods
M05. Information about the instrument, calibration
M06. Quality assurance procedures
M07. Information about limitations of the dataset, information that ensures correct interpretation of the dataset
M08. People involved in the creation or processing of the dataset

DATA ACCESS AND SHARING
A01. Recommended citation for the dataset
A02. License information, restrictions on use
A03. Confidentiality information

DATA SPECIFIC INFORMATION (ABOUT THE DATA THEMSELVES)
D01. Full names and definitions for columns and rows
D02. Explanation of abbreviations
D03. Units of measurement
D04. Symbols for missing data

RELATIONSHIPS
R01. Publications based on this dataset
R02. This dataset derives from… (other dataset)
R03. This dataset is related to… (documents, dataset)
R04. References of publications used to create the datasets

Additional information can be included, depending on the needs of the project for future use.
Furthermore, Excel codebooks will be created concerning the surveys and the main datasets, explaining all the variables used (includes e.g., variable names and labels, origin of the scales that are used, link between the variables and the survey questions, etc.). Also the informed consent forms of the studies will be retained and stored. Additionally, metadata will also be captured via the KU Leuven research data repository (see below): RDR. Uploading datasets in RDR requires a systematic overview of basic metadata like title, author, contact information, key words etc. (and possibility to add extra metadata like grant or geographical information). The metadata in RDR is always publicly available, even if the data files themselves are restricted or closed.

Will a metadata standard be used to make it easier to find and reuse the data ? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse.
A metadata standard is automatically applied when depositing data in a trusted data repository. As the PhD researcher will deposit the project data in the KU Leuven RDR, the metadata will be created according to the “Metadata Model” of KU Leuven. With the

metadata categories mentioned in the README-file and the RDR, the 15 principal elements of the Dublin Core Standard are also expected to be covered (Creator, Contributor, Publisher, Title, Date, Language, Format, Subject, Description, Identifier, Relation, Source, Type, Coverage, and Rights). Were any element to be missing, the PhD researcher will add it to the metadata.

Data Storage & Back-up during the Research Project

Where will the data be stored?
The personal data of the entire project will be saved on the researcher's personal and secured KU Leuven OneDrive for Business drive. After pseudonymization, the non-pseudonymized data will be deleted from the drive. The researcher will grant access to these data through OneDrive to the project's promoter and supervisor – Elfi Baillien - as to collaborate on the different studies. Paper data will be stored by the PhD researcher in a locked drawer or cupboard that can only be accessed by him/herself. After the project has been finished, all paper data will be handed over to the supervisor who will store these data in their office in a locked drawer or cupboard that can only be accessed by him/herself.

How will the data be backed up?
The OneDrive for Business folders of KU Leuven have automatic, cloud-based daily back-up procedures.

Is there currently sufficient storage & backup capacity during the project? If no or insufficient storage or backup capacities are available, explain how this will be taken care of.
Yes. We consider that there will be sufficient storage and backup capacity during the project. The standard offer of OneDrive for Business is 2 TB but can be extended to 5 TB upon motivated request.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?
The personal data of the entire project will be saved on the researcher's personal and secured KU Leuven OneDrive for Business drive. Personal data will be deleted as soon as possible, that is, after the coding process or after prizes/ feedback reports have been distributed among respondents. In other words, after pseudonymization, the non-pseudonymized data will be deleted from the drive. In addition, the access to this secured drive requires a password. These measures should prevent the risk of lost or stolen data when laptops are lost. Regarding access to the files, the doctoral researcher will grant promoter and supervisor Elfi Baillien access to all the files during the project. Since master's students may be involved in some of the studies (i.e., master's thesis), they will have access to (only for their research questions) relevant data as well. However, only pseudonymized data (password protected) will be shared with them through the OneDrive. We will have them sign a form stating that all data must be deleted after completion of the analyses.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?
No extra costs are expected. OneDrive is part of the Microsoft 365 Education A3 plan. The cost of the Microsoft 365 Education A3 licenses within the EES agreement is financed centrally for all KU Leuven students and the majority of active KU Leuven staff.

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).
Personal data will be deleted as soon as possible, that is, after the coding process or after prizes/feedback reports have been distributed among respondents. The remaining, coded data will be retained for 10 years and deleted after, as recommended by KU Leuven's guidelines on research data management. Recordings (interviews in study 1) will be deleted after transcription/coding of the data.

Where will these data be archived (stored and curated for the long-term)?
The personal data of the entire project will be saved on the researcher's personal and secured KU Leuven OneDrive. After pseudonymization, the non-pseudonymized data will be deleted from the drive. After the project has been finalized, all the data remaining on the PhD researcher's OneDrive will be permanently transferred and stored on KU Leuven's OneDrive central storage facilities of the research unit Work and Organisation Studies, which is secured and can only be accessed by the researchers of the research group. The data will be retained for 10 years, as recommended by KU Leuven's guidelines on research data management. This drive is secured and is backed-up automatically on the cloud. Only personnel from the research group has access to this drive via a password. Additionally, the data will also remain available in the KU Leuven RDR (see below) for a minimum of 10 years.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?
No additional costs will be necessary to preserve data on the secured drive for a period of 10 years.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available.
Yes, as restricted data (upon approval, or institutional access only).

Whenever the PhD researcher completes a research paper (i.e., whenever a work package is finished) the related data will be made available. This data will be deposited in the KU Leuven institutional research data repository, the RDR. The data will be filed under restricted access rights. This means that files are restricted in access but do allow for access requests to be made. In light of a collaboration, the data will thus be shared with third party researchers upon requests (based on clear research questions and hypotheses). The data will also be shared with third parties that the PhD researcher and promotor themselves involve in the project or whenever the sharing of data is necessary for the successful completion of the project. After the PhD project finishes, the data will still be accessible upon request in view of collaborations and/or questions on the data regarding possible meta-analyses.

The above will hold for all survey data. The interview data from the qualitative study 1 will be filed under closed access rights, leaving no option to request access to the files. This is done in light of privacy and ethical considerations since the complete pseudonymization of the respondents is more difficult to guarantee.

If access is restricted, please specify who will be able to access the data and under what conditions.
As mentioned above, the interview data of study 1 will be closed access. However, the access to survey data is restricted to sharing upon request based on clear research questions and hypotheses. Given the nature of the request, the PhD researcher and/or promotor will decide to grant access to the data or not.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.
Yes, privacy and ethical considerations prevent the qualitative interview data from being shared.

Where will the data be made available? If already known, please provide a repository per dataset or data type.
The data will be made available in the KU Leuven institutional research data repository, the RDR.

When will the data be made available?
Upon publication of research results: whenever a research paper is completed, the related data will be made available as part of the publication of the research.

Which data usage licenses are you going to provide? If none, please explain why.
In line with KU Leuven guidelines, the restricted and closed data make it impossible to legally set a license as there is no open data available for a license to be applied to. As a result, we will indicate the "custom KU Leuven" license selection when uploading the data in RDR. Whenever a request is made to access the data, a data transfer/sharing agreement will be drawn up in which the terms of use will be agreed upon. The RDM support desk will provide guidance in the creation of these data transfer agreements.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.
Once a dataset is created in the KU Leuven RDR, the RDR automatically assigns a DOI to permanently identify it. This DOI is reserved for the dataset and will be activated upon publication.

What are the expected costs for data sharing? How will these costs be covered?
None, as the university's own research data repository services will be used, this is free of charge.

Responsibilities

Who will manage data documentation and metadata during the research project?
The PhD researcher (Charlotte Franckx) and promotor (Elfi Baillien).

Who will manage data storage and backup during the research project?
The PhD researcher (Charlotte Franckx) and promotor (Elfi Baillien) will be responsible for all data storage.

Who will manage data preservation and sharing?
The PhD researcher (Charlotte Franckx) and promotor (Elfi Baillien).

Who will update and implement this DMP?
The PhD researcher (Charlotte Franckx) and promotor (Elfi Baillien).