Closed-loop neuro-steered beamforming for hearing aid applications

FWO Data Management Plan

ADMIN DETAILS

Project Name: Closed-loop neuro-steered beamforming for hearing aid applications
Grant Title: 1S14922N
Principal Investigator / Researcher: Iustina Rotaru
Institution: KU Leuven

1. GENERAL INFORMATION

Name applicant
Iustina Rotaru

FWO Project Number & Title
1S14922N
Closed-loop neuro-steered beamforming for hearing aid applications

Affiliation

- KU Leuven

2. DATA DESCRIPTION

Will you generate/collect new data and/or make use of existing data?

- I will make use of existing data and I will also generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project.

Type of data	Format	Volume	Description
Scanned images of general questionnaire	.pdf	2,5 MB per participant (6 pages)	The participant completes the questionnaire on paper. It will be scanned afterwards.

Scanned images of laterality preference questionnaire	.pdf	1,2 MB per participant (2 pages)	The participant completes the questionnaire on paper. It will be scanned afterwards.
Scanned images of the hearing threshold. The numbers will also be written down in an csv file.	.pdf .csv	500 KB per participant (1 page)	Tone audiometry. Beep tones of different frequencies and loudness are presented via a headphone to the participant. The participant says 'yes' when he hears the tone and this threshold will be written down by the researcher. This paper with the thresholds will be scanned.
Scanned images of speech reception threshold. Numbers will also be written down in a csv file.	.apx .pdf .csv	500 KB per participant (1 page)	Speech audiometry test. Matrix sentences will be presented to the subject via insert-phones. The stimuli are sent via the software program Apex or Python. Apex generates a results file.
EEG (electro-encephalo-graphy) data	.apr .bdf	+/- 15 GB per participant	These are the core data of our experiments, which consist of maximum 10 conditions where electroencephalography is being measured via scalp electrodes.
Miscellaneous experimental files	.py .apx .apf .js .wav .mp4 .json .jpeg .png	~ 2,5 GB (experiment files + audio and visual stimuli per experiment)	The experiment files contain the folder structure of the experiment, metadata for the different participants, the audio and visual stimuli, as well as the randomization order of the experimental conditions.
Eye tracking data	.tsv	+/- 500 MB per participant	We'll use an eye-tracker device (Gazepoint GP3 HD 150HZ) to track eye movements while the study participants listen to audio stimuli. Operating the eye tracker is done via Python and Gazepoint hardware. During a measurement, .tsv files are generated - they contain the gaze direction of each eye in screen-coordinates.
Analysis files	.m .mat .npy .csv	15 GB per year	Matlab, Python and csv files used to pre-process the stimuli, analyse the data and generate intermediate result files.

R files	.r .csv	1 GB per year	Files from R software to statistically analyse the data.
Administrative files	.pptx .md .pdf .docx	50 GB per year	Files that describe the dataset, literature, methods, etc.

3. LEGAL AND ETHICAL ISSUES

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- Yes

Privacy Registry Reference:
The study is being approved by the Ethic Committee *Ethische Commissie Onderzoek UZ/KU Leuven* (S57102).
Short description of the kind of personal data that will be used:
Some basic information is gathered from the subjects (name, age, gender, mother tongue, cochlear implant or not, …) though strictly limited to the absolute necessary data. Furthermore, the collected data are pseudonymised.

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, add the reference to the formal approval by the relevant ethical review committee(s)

- Yes

The project involves experiments on Human subjects and therefore is being approved by the Ethic Committee *Ethische Commissie Onderzoek UZ/KU Leuven* (S57102).

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?
Not at this moment. In case IP possibilities will occur, we will contact LRD.

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. DOCUMENTATION AND METADATA

What documentation will be provided to enable reuse of the data collected/generated in this project?

New lab members get a training on how to perform the experiments how and where to store and handle the collected data. Furthermore, standard guidelines and instructions are available as .pdf and/or .readme, stored on network drive. They contain best practices regarding the practical side of the experiments (set-up, parameters, …) as well as policies about how to treat the subjects, how to handle and where to keep sensitive information, etc. In this way, the information given during the training can easily be re-read, refreshed.

This, in combination with the fact that the original collected data, are stored on backed-up drives (Microsoft OneDrive and two external HDDs), in a standard format, should make it possible to understand and reuse the data. The data will be stored based on the BIDS structure (see below).

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- Yes

The data will be stored based on the BIDS structure, a standard for organizing neuroimaging and behavioural data in a specific folder hierarchy (it is described in more detail at bids.neuroimaging.io). This standard will be used to organize the data so that it matches the format of other datasets in our field, thus facilitating reuse and reproducibility.

5. DATA STORAGE AND BACKUP DURING THE FWO PROJECT

Where will the data be stored?

The data will be stored on KU Leuven administered drives (large volume storage drive and OneDrive). In order to be able to analyse the data, some files will need (temporarily) to be stored on the encrypted PC hard drive (because calculations from a non-local source are too slow and lead to computational failures). Once analysed, the raw data are again removed from the local hard drive.

How is backup of the data provided?

Since the data are stored on KU Leuven storage, the general ICT back-up Policy is applied.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

Since the data are stored on KU Leuven servers, and these drives are expandable in blocks, the backup capacity is technically not an issue.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

The estimated cost for data storage and back-up over the 4-year FWO project is around 700 Euro and will be paid from the project's bench fee ('werkingskosten').

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?
The lab policy ensures that the researchers have access only to the data from the project they are involved in. Furthermore, the data for longer-term storage are kept on separate drives with a) limited access (only a limited set of people have access) and b) an overwrite and delete protection (based on read-write access) in order to prevent accidental loss of these data.

Prof. Francart is the only person who has access to the key information for the identification of the subjects. His back-up (only to be used in extreme case) is Jan Wouters and Astrid van Wieringen (both PI's within the lab).

6. DATA PRESERVATION AFTER THE FWO PROJECT

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).
In accordance with the applicable law, the data of this study will be retained for 20 years (KB 30 juni 2004). These terms (and conditions) are also mentioned in the Informed Consent that will be signed by the subject, prior to their participation.

Where will the data be archived (= stored for the longer term)?
The data are stored on a KU Leuven, ICT-managed Large Volume Storage drives especially designed for archive storage. Personal data will be kept in a similarly secure way. However, these data will be stored at a different location (different server) with other (restricted) access rights (see also above).

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?
The estimated cost for the 5 year retention period is around 1000 Euro.
Given the way projects are financed at KU Leuven, we cannot predict whether funding will be available after the current project ends, unless we are allowed to set aside funding from the project for later use.

7. DATA SHARING AND REUSE

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- Yes. Specify:

To be decided.

The key to identify the subject from which the data were obtained, is definitely not sharable (without the consent of the subject) and/or this will also depend on IP issues.

Which data will be made available after the end of the project?
To be decided, this will depend on IP issues and/or on the consent of the participant.

Where/how will the data be made available for reuse?

- Other (specify):

Not yet applicable, this will also depend on the choice the participant opted for (e.g. direct sharing with other research institution vs sharing on public accessible database).

When will the data be made available?
Not yet applicable

Who will be able to access the data and under what conditions?
Only the limited set of researchers working on this specific project will have access to the data. Furthermore, the data are pseudonymised, so even for the researchers, there is no (recognisable) link between the data and the test person. Only a very limited number of authorized persons (1-2) have access to this identifier-key. In case sharing of the data within the lab is necessary/favourable, the PI will decide which data will be shared. For sharing outside the lab, the consent of the subject will be asked.

What are the expected costs for data sharing? How will the costs be covered?
Not yet applicable

8. RESPONSIBILITIES

Who will be responsible for data documentation & metadata?
The researchers themselves have been instructed on how to collect and document the data.

Who will be responsible for data storage & back up during the project?
The researchers themselves have been instructed on how to handle and store the data. Back-up of the data is arranged by ICTS- KU Leuven (back-up policy dependent on the kind of storage).

Who will be responsible for ensuring data preservation and reuse?
The PI through highly restricted overwrite and delete authorisation. It is also the PI's decision which data are shared with other co-workers and how this will be arranged.

Who bears the end responsibility for updating & implementing this DMP?
The PI, Prof. Tom Francart bears the end responsibility of updating & implementing this DMP.