FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	Brent Sterckx (0000-0001-8907-5532)
Contributor name(s) (+ ORCID) & roles	Prof. Jan Wouters (0000-0002-0093-698X): supervisor Prof. Marc Moonen (0000-0003-4461-0073): co-supervisor Dr. Robin Gransier (0000-0001-7429-6439): co-supervisor
Project number [1] & title	Personalised stimulation strategies for cochlear implants using EEG source analysis
Funder(s) GrantID [2]	1SHI924N
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	A cochlear implant (CI) is a bionic neuroprosthesis that restores auditory perception in individuals with severe hearing loss by directly stimulating the auditory nerve via intracochlear electrodes. Despite the CI's success, speech perception with a CI is still far worse compared to normal-hearing listeners, and highly variable across CI users, especially in everyday adverse listening conditions. For speech perception, temporal modulations in the speech envelope (TEMs) are crucial, and different cortical hemispheres specialise in processing TEMs at different time-scales. For speech perception with a CI, it is crucial that the TEMs are properly encoded in the auditory periphery, but equally, that the auditory cortex can process them correctly. However, evidence shows that hemispheric activation is altered in CI users, but insight in the hemispheric processing of TEMs is currently impossible due to CI stimulation artifacts obscuring the neural responses. Yet, a hemispheric specialisation-activation mismatch for certain time-scales can be present for which currently stimulation strategies do not account, and that decreases performance. This project develops a novel EEG-based source analysis method robust to CI stimulation artifacts, to study the individual hemispheric activation patterns in CI users. The insights will then be used to develop personalised, neuro-inspired stimulation strategies that compensate for specialisation-activation mismatches to improve speech perception.

2. 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 (add rows for each dataset you want to describe).

Two datasets will be collected:

- 'ASSR' – This dataset will encompass the data collected for both WP1 and WP2. It will primarily consist (see below for a detailed description of the contents of this dataset) of auditory steady-state responses (ASSRs) and speech performance results recorded from both normal-hearing (NH) individuals as well as cochlear implant (CI) users.
- 'iStrat' – This dataset will encompass the data collected in WP3. It will contain ASSRs, software to generate personalised stimulation strategies from these ASSRs, the generated or otherwise tested strategies, and the speech performance results for with those strategies.

Each datasets will also contain the stimuli used in the experiments, as well as relevant personal details about each participant (e.g. sex, age, hearing status, etiology, …). Additionally with each dataset, software will be developed to analyse the data.

				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
ASSR iStrat	(scans of) personal info/background questionnaires	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	PDF	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	A4 binder
ASSR iStrat	(scans of) Informed Consent Forms	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical	PDF	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB	A4 binder

				☒ Textual ☐ Model ☐ Software ☐ Other:		☐ < 5 TB ☐ > 5 TB ☐ NA	
ASSR iStrat	(behavioural) Audiological (test) data	☒ Generate new data ☒ Reuse existing data (if previously performed tests can be obtained, e.g. through consent to obtain from individual'	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	PDF CSV CDX (Cochlear® CustomSound Pro) JSON XML	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	A4 binder
ASSR iStrat	EEG Recordings	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	BDF	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA Estimates: ~3 GB/condition ~30 GB/session ~1 TB total	
ASSR iStrat	3D EEG electrode positions	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model	TSV	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB	

				☐ Software ☐ Other:		☐ NA	
ASSR iStrat	Experiment files/stimuli (files to set up and run the experiments)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☒ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☒ Other: Stimulation software specific binary files	APX RBADATA ASEQ MAT WAV XML CDX	☐ < 1 GB ☐ < 10 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
ASSR iStrat	Behavioural test results	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images (scans) ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	PDF CSV JSON XML APR	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	A4 binder
ASSR iStrat	Analysis scripts	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☒ Software ☐ Other:	.m .py .R	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

ASSR iStrat	Processed data (i.e. after pre-processing and response extraction)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.mat CSV	☐ < 1 GB ☐ < 10 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
ASSR iStrat	Results (e.g. statistics and figures)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.mat CSV .eps .pdf .svg .png .tiff	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
ASSR iStrat	Information about protocol, study aims, methods, datasets, literature,… (or scans of these files)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.pdf .docx CSV .txt .xlsx .pptx	☐ < 1 GB ☒ < 10 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

GUIDANCE: *The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.* *RDM Guidance on data*	
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.	NA
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: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: This project (all datasets) involves experiments on humans which are subject to approval by the competent ethics committee, Ethics Committee Research UZ/KU Leuven. An application is being prepared.

Will you process personal data[3]? 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) ☐ No Additional information: For both datasets, personal information (such as name, age, gender, and address) as well as medical background information (only that pertaining to audiological and neurological background and which could influence the results, such as information about previous hearing interventions/treatments or previous brain injury) will be captured in the questionnaire and processed. Moreover, EEG data and results from behavioral tests (e.g. speech in noise tests) are related to the subject's health and are thus special category data according to the GDPR and require careful ethical consideration during processing. Personal information will also be captured in the ICF, as required by both applicable law and GCP standards. An application for approval by the competent ethics committee is currently being prepared and will include a PRivacy and EThics (PRET) review which also registers the processing of personal data in KU Leuven's "internal processing of personal data register".
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 data type where appropriate.	☐ Yes ☒ No If yes, please comment: Not at this moment. In case IP possibilities arise, we will contact the KU Leuven Research and Development Department to advise on necessary steps.
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:

[3] See Glossary Flemish Standard Data Management Plan

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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. 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). *RDM guidance on documentation and metadata*.	New lab members will receive training on how to perform the experiments. They will also be informed about how and where to store and handle the collected data. Furthermore, standard guidelines and instructions are available as .pdf, .docx and/or .readme, stored on network drives. They contain best practices regarding the practical side of experiments as well as policies about how to treat the participants, how to handle and store sensitive information. The data will be stored according to the standard BIDS format, which includes metadata where appropriate. Despite being structured according to the standard BIDS format, the data is not easily shared across the lab. Personal information about the participants is restricted and only accessible by the PI. The PI also has the final decision about sharing the data, which is constrained by the protocol approved by the ethical committee and by the informed consent signed by the participant.

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. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: The data will be stored based on the existing BIDS structure, which is already used for neuroimaging and behavioral data (see also website bids.neuroimaging.io). We will use this data structure in order to have a uniform way of data arrangement across studies. If no, please specify (where appropriate per dataset or data type) which metadata will be created:

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☐ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premises ☒ Large Volume Storage ☐ Digital Vault ☒ Other: Data will be stored on encrypted storage drives with restricted physical access to ensure that only the PI, co-investigators, and persons approved by the PI will have access. To facilitate analysis, some files will (temporarily) need to be stored on encrypted portable storage drives or laptops. Once the data is analyzed, these data are again removed from the these media.

How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☐ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☒ Other (specify) All data will be stored on at least two different encrypted storage drives to back up each other. Access to these drives is restricted to ensure that only the PI and co-investigators will have access. A master document is used to keep track of the location of each dataset and the number of copies.
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 ☐ No If no, please specify: As can be seen from the size estimates in the table above in Section 1, regular storage drives are available with sufficient capacity to store and backup the datasets.
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	To prevent unauthorized access, data is stored on encrypted storage drives with restricted physical access to ensure that only the PI, co-investigators, and persons approved by the PI will have access. To prevent data loss, all data has at least one back-up and physical access to these drives is restricted to the PI and co-investigators. To prevent accidental data loss, overwrite and delete protection is used. As an added security measure, all data will be pseudonymized using a subject-specific identifier that does not allow to identify the subject. Personal data not needed for analysis (but needed for administrative and legal purposes, such as name, address, date of birth…) as well as the coding table linking subjects to identifiers, will be stored physically separate from the pseudonymized data with the same abovementioned security measures in place. Only the PI has access to the identifiable data, as well as the co-investigators on a need-to-know basis and upon approval by the PI. Transferring data will only be done using secure, encrypted methods, such as with encrypted hard drives or using Belnet filetransfer with data encryption enabled.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The estimated cost for data storage and back-up is around 700 Euros and will be covered by the allotted bench fee.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five 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...). *Guidance on data preservation*	☐ All data will be preserved for 10 years according to KU Leuven RDM policy ☒ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain) Additionally, participants agree to these terms and conditions through their written informed consent.
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☒ Other (specifiy): The storage drives used for storage can be used to preserve the data during the retention period.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The storage drives used for storage can be used to preserve the data during the retention period.

6. 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. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☒ Yes, as restricted data (upon approval, or institutional access only) ☒ No (closed access) ☐ Other, please specify: Due to the sensitive nature of the data captured (personal and health information), these cannot be shared openly due to both legal and ethical aspects. The following restrictions are in place: • *All pseudonymized data*: access restricted, requiring approval by the PI and only within the constraints of the subjects' consent and the protocol approved by the competent ethics committee. Access will be governed by an agreement (e.g. a data transfer agreement). • *Directly identifiable personal data* (e.g. name, address, date of birth) and *coding file*: closed access To protect potential IP, the following restrictions are in place: • *Software, experiment files and stimuli*: access restricted, requiring approval by the PI. The PI will decide which files (if any) will be shared and under which conditions or license (if any).
If access is restricted, please specify who will be able to access the data and under what conditions.	Only the PI can approve the sharing of data and under which conditions (e.g. a DTA), and only within the constraints of the subjects' consents and the protocol approval by the competent ethics committee.
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 aspects ☒ Yes, intellectual property rights ☒ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No If yes, please specify: (see above)

Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ KU Leuven RDR ☐ Other data repository (specify) ☒ Other (specify) The personal data is stored in a restricted database of the lab only accessible by the PI, and by the co-investigators on a need-to-know basis and upon approval by the PI. Experiment data are pseudonymized and stored only in an encrypted storage medium as outlined above, for maximal data security.
When will the data be made available?	☐ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify) To be decided upon.
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☐ CC-BY 4.0 (data) ☒ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☐ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☒ No

What are the expected costs for data sharing? How will these costs be covered?	None. NA

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The researcher himself has been instructed on how to collect and document the data.
Who will manage data storage and backup during the research project?	The researchers himself has been instructed on how to handle and store the data. Back-up of the data is managed by the PI.
Who will manage data preservation and sharing?	The PI through highly restricted overwrite and delete protection. The PI will also decide which data are shared outside of the project and under which conditions, within the constraints of the subjects' consent and the protocol approved by the competent ethics committee.
Who will update and implement this DMP?	The researcher themselves carries some responsibility. The PI bears the end responsibility of updating and implementing this DMP.