FWO DMP Template - Flemish Standard Data Management Plan

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	**ARNAU SANS DUBLANC; https://orcid.org/0000-0002-6132-8187**
Contributor name(s) (+ ORCID) & roles	**KARL FARROW, promotor; https://orcid.org/0000-0003-1409-096X** **MAXIMILIAN JOESCH, Co-promotor; https://orcid.org/0000-0002-3937-1330**
Project number[1] & title	Dissecting visual selective attention: from brain-wide networks to collicular cell-types.
Funder(s) GrantID[2]	12ZZZ23N
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel X Other: Neuroelectronics research flanders; Institute of Science and Technology Austria (ISTA); Provide ROR[3] identifier when possible:

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. 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.

[3] Research Organization Registry Community. https://ror.org/

Please provide a short project description	At any given time, we attend to only a small fraction of the available sensory information. This process, known as selective attention, allows us to effectively interact with behaviorally relevant features of our environment. Selective attention can be driven by both goal-directed "top-down" and sensory-driven "bottom-up" processes. Although their interaction instructs what we attend to, a global perspective of the pathways and mechanisms involved in their coordination is still largely unknown. Here, I will determine the neural dynamics associated with the weighting of bottom-up and top-down attention to guide behavior. To do this, I will first unbiasedly define the set of brain areas activated during a task that requires top-down visual selective attention using functional ultrasound imaging of the entire brain. Second, I will assess the impact of presenting innately salient stimuli during the same task. Third, I will manipulate a subset of the identified circuits to test their causal role. Finally, to gain insights into the mechanisms underlying the observed brain-wide changes, I will determine, using multi-photon imaging, how different cell-types integrate attentional information. Here, I will focus on the superior colliculus, a midbrain area known to be a central piece of the puzzle. Together, these experiments will provide fundamental insights into how the brain coordinates internal and external drivers of attention that guide behavior.

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[4].

				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
vfUSI	Volumetric functional ultrasound whole brain imaging data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☒ Experimental	.mat	< 10 TB	
Face camera	tracking of animal's eye and face	☒ Generate new data	☒ Digital	☒ Experimental	.avi .tiff	< 5 TB	
Ball tracking	Tracking of motion of the animal	☒ Generate new data	☒ Digital	☒ Experimental	.csv	< 1 GB	
Lick sensor	Tracking of the task-related responses of the mice	☒ Generate new data	☒ Digital	☒ Experimental	.csv	< 1 GB	
2p	Two-photon calcium imaging of SC activity	☒ Generate new data	☒ Digital	☒ Experimental	.tiff	< 5 TB	

[4] Add rows for each dataset you want to describe.

Histology	Histology of vfUSI and 2p experiments	☒ Generate new data	☒ Digital	☒ Experimental	.tiff	< 100 GB	
Manuscripts	Manuscripts resulting from project	☒ Generate new data	☒ Digital	☐ Observational ☐ Experimental ☒ Compiled/ aggregated data ☐ Simulation data ☐ Software ☒ Other ☐ NA	.docx	< 1 GB	
Event log	Event logging of visual stimuli	☒ Generate new data	☒ Digital	☒ Experimental	.csv .STIMlog (text file)	< 1 GB	

Guidance: *Data can be digital or physical (for example biobank, biological samples, …). Data type: Data are often grouped by type (observational, experimental etc.), format and/or collection/generation method.* *Examples of data types: observational (e.g. survey results, sensor readings, sensory observations); experimental (e.g. microscopy, spectroscopy, chromatograms, gene sequences); compiled/aggregated data[5] (e.g. text & data mining, derived variables, 3D modelling); simulation data (e.g. climate models); software, etc.* *Examples of data formats: tabular data (.por,. spss, structured text or mark-up file XML, .tab, .csv), textual data (.rtf, .xml, .txt), geospatial data (.dwg,. GML, ..), image data, audio data, video data, documentation & computational script.* *digital data volume: Please estimate the upper limit of the volume of the data per dataset or data type.* *physical volume: Please estimate the physical volume of the research materials (for example the number of relevant biological samples that need to be stored and preserved during the project and/or after).*	
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.	No reuse of data

[5] These data are generated by combining multiple existing datasets.

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, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☒ Yes, animal data ☐ Yes, dual use ☐ No If yes, please describe: The entirety of this project will be based on animal experimentation (mice). In order to generate any of the data types of this project, mice will undergo one or two surgeries to allow head-fixation, brain imaging (craniotomies), neural circuit manipulations, and behavioral training on a spherical treadmill. Additionally, mild water deprivation will be used for behavioral training. All experiments are done under the approval of the ethical committee of the KU Leuven and the ethical committee of ISTA for the experiments that will be performed there (Calcium imaging data). Any distress caused by the surgeries or training is short-lived (1-5 days for recovery) and does not cause permanent health issues.
Will you process personal data[6]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:
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:

[6] See Glossary Flemish Standard Data Management Plan

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:
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).	During the project, data will be stored on the labs file server (capacity for this project: 50 TB) for easy access of the data. Manuscripts are stored on a group-internal OneDrive that is based on the imec corporate account. After each experiment, the data is saved to the labs file server and analyzed there. After analysis, all data is stored at the institute-internal file server (capacity for this project: of 50 TB NERF, 105 TB IST). Each animal will be assigned an experiment number and data folder that will be duplicated into the labs servers e.g,: //NERFFS17/farrowwip (for analysis) and //NERFHF01/farrowarch (for permanent storage) at NERF or fs.ista.ac.at/archive (storage) and fs.ista.ac.at/group/joeschgrp (analysis), at ISTA. In each of these folders a metadata file (.txt) will be created stating the type of experiments performed to the mouse (behavior, brain imaging...), gender, mouse line, fluorescent reporters present in the mouse brain, cage number, date of birth, ear tag. Inside each experiment folder, a common sub folder structure will be followed including; folder with type of data (imaging, histology, behavior...)/folder named with the date when the data was generated (session folder)/folder with raw data or pre-processed or analysed data. Inside each session folder a .txt file will be saved with details of that day's experiment. E.g: general quality of the experiment, relevant events particular of that experiment, number of trials, type of session (training level 1, 2, 3, experimental session), task used in session (type 1, 2, 3), observational comments on the state of the animal during the experiment. Additionally, each stage of each mouse experiment will be logged on an institutional outlook OneNote Labbook. Here each procedure performed to the mouse (surgery, habituation session, training session, experimental session), and the details of the mouse (gender, date of birth, mouse line...) will be described in detail. Also, all mice, when receiving an experiment number will be logged in a excel file stored in the institutional Onedrive of the lab, where all mice that have been used in the lab are logged with experiment number, user, date of first intervention, type of experiment, cage card, mouse line, if virus is injected, what virus, short experiment description, ear tag, gender, general observation/comments. Finally, in order to easily find the data relevant to this project, in the institutional Onedrive of the lab, we have folders that are specific of each project. In the folder specific to this project an excel file called "ExpDB" will be created. Inside this file all experiment numbers associated with this project will be reported, type of experiments, mouse line, types of data types available from each session, specific location of the data (e.g. in which partition of NERFFS server), and stage of analysis.

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: If no, please specify (where appropriate per dataset or data type) which metadata will be created: In project folder: -Metadata file will be created for the project to record: All experimental numbers/mice included in the project, type of experiments, which subproject they belong to, task used, data types per session, mouse line, analysis stage, location of the data. In mouse folder: -Metadata file will be created for each mouse to record: Type of experiments performed to the mouse (behavior, brain imaging...), gender, mouse line, fluorescent reporters present in the mouse brain, cage number, date of birth, ear tag. In session folder: -Metadata file will be created for each session to record: General quality of the experiment, relevant events particular of that experiment, number of trials, type of session (training level 1, 2, 3, experimental session), task used in session (type 1, 2, 3), observational comments on the state of the animal during the experiment.

4. Data Storage & Back-up during the Research Project

Where will the data be stored?	**Data will be stored in the institutional servers of NERF (IMEC) and ISTA.** **All data generated in the farrow lab will be duplicated and stored in the archiving server (NERFHF01) and the working server (//NEFFS17/farrowlabwip2023/Data) for data generated at ISTA: fs.ista.ac.at/archive (storage) and fs.ista.ac.at/group/joeschgrp (analysis)**
How will the data be backed up? *What storage and backup procedures will be in place to prevent data loss? Describe the locations, storage media and procedures that will be used for storing and backing up digital and non-digital data during research.*[7] *Refer to institution-specific policies regarding backup procedures when appropriate.*	**All data generated in the farrow lab will be duplicated and stored in the archiving server (NERFHF01) and the working server (//NEFFS17/farrowlabwip2023/Data), for data generated at ISTA: fs.ista.ac.at/archive (storage) and fs.ista.ac.at/group/joeschgrp (analysis)**
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 yes, please specify concisely: The capacity of the file server for this project is 50 TB NERF, 105 TB ISTA for easy access of the data, which is well beyond the data storage space that is expected to be needed. Manuscripts are stored on a group-internal OneDrive that is based on the imec corporate account. If no, please specify:

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

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.* [7]	All computers at the institute are password protected. Additionally, to access the servers where the data is stored one can only do it through a password protected computer or by connecting directly to the server, which is also password protected. Only members of the lab receive access to the computers and servers of the lab.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	cost is 28€/TB/year for storage and back up in a second server. All costs related to the storage of data are taken care through grants and institutional funding from NERF.

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...).	**All data obtained from this project will be retained for at least 5 years after the end of the project.**
Where will these data be archived (stored and curated for the long-term)?	**All data generated in the farrow lab will be duplicated and stored in the archiving server (NERFHF01) for long-term storage.** **Additionally, all data related to published results derived from this project will be stored in the KU Leuven research data repository for public access.**

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	**KU Leuven Research Data Repository allows free storage of up to 50GB. Since the data that will be uploaded will be pre-processed data we do not expect to exceed that amount.** **Regarding all raw data produced and stored during the project, cost is 28€/TB/year for storage and back up in a second server. Costs are paid through grants by the lab.** **All costs related to the storage of raw data stored are taken care through grants and institutional funding from NERF.**

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, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☐ Other, please specify: All the data generated in this project will be made accessible to third parties after publication of the results, through the institutional research data repository or the KU Leuven.
If access is restricted, please specify who will be able to access the data and under what conditions.	Only lab members and collaborators in this project (e.g., the Urban lab and the Joesch lab) will have direct access to the data generated in this project in order to help with the analysis and manuscript writting.
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:

Where will the data be made available? If already known, please provide a repository per dataset or data type.	**All data generated in the farrow lab will be duplicated and stored in the archiving server (NERFHF01) for long-term storage and reuse upon reasonable request. All published results will be made accessible in pre-processed format corresponding to the data from each published figure, through the institutional research data repository of the KU Leuven.**
When will the data be made available? *This could be a specific date (dd/mm/yyyy) or an indication such as 'upon publication of research results'.*	**UPON PUBLICATION OF RESEARCH RESULTS**
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.* *Example Answer: E.g. "Data from the project that can be shared will be made available under a Creative Commons Attribution license (CC-BY 4.0), so that users have to give credit to the original data creators."* [8]	All published data will be made available under a CREATIVE COMMONS ATTRIBUTION LICENSE CC-BY-NC-4.0 and users will be expected to give credit to the creators of the data through citation of the original published work.

[8] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

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 ☐ No If yes: **Not available yet.**
What are the expected costs for data sharing? How will these costs be covered?	**Sharing data has no cost within NERF. NERF uses Globus for sharing data, for which NERF uses the Vlaams Supercomputer Centrum subscription.** **Also, KU Leuven Research Data Repository allows free storage of up to 50GB. Since the data that will be uploaded will be pre-processed data we do not expect to exceed that amount.**

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	**The fellowship holder**
Who will manage data storage and backup during the research project?	**The fellowship holder and promotors Karl Farrow and Max Joesch**
Who will manage data preservation and sharing?	**Karl Farrow and Max Joesch**
Who will update and implement this DMP?	**The fellowship holder**