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	**Marie Van Dijck 0009-0009-7979-6390**
Contributor name(s) (+ ORCID) & roles	**Promotor: Prof. Dr. Eve Seuntjens 0000-0002-0126-461X**
Project number [1] & title	1181025N: Ontogeny of cell types and neural activity in the Octopus vulgaris optic lobe
Funder(s) GrantID [2]	1181025N
Affiliation(s)	X 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	Octopuses have evolved large nervous systems and camera-type eyes like humans, and offer a unique model for studying evolution of vision. In coleoid cephalopods, visual information is mainly processed in the optic lobe, which contains an outer laminated cortex that resembles the vertebrate retina, and inner medulla. Our lab has shown that already at hatching, the O. vulgaris optic lobes (OL) contain a diversity of cell types. These are generated from a temporary embryonic structure called the lateral lip, which is spatially patterned to generate neurons for the optic lobe and other brain regions. How each optic lobe cell type gets specified over time and how it wires up remains a mystery. This project will unravel the ontogeny of optic lobe cell types and circuits. First, I will reveal spatial and temporal patterns of OL neurogenesis using lineage tracing and HCR co-staining, providing a fate map for OL cell types. Second, I will identify transcription factors that specify OL cell types by snRNAseq at different time points during development. Third, I will connect molecular identities to morphological and functional phenotypes at hatching in the OL cortex using Ca imaging, patch-clamp electrophysiology and RNA sequencing. Collectively, this will offer deep insight into the ontogeny and functionality of the octopus early visual circuitry. As such, these data will serve as a model dataset for understanding the development and function of evolutionary expanded neural systems.

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 [3].

				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
Imaging data	Confocal images; Lightsheet images; fluobino images; obtained from HCR, IHC, injections, calcium imaging and/or patch-clamp from sections or wholemounts of the optic lobe.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	tif, .jpg, .jpeg, .png, .czi, sis, .mp4, .avi, .ai, .pdf, .pptx	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ > 5 TB ☐ NA	
Sequencing data	Single nuclei RNA sequencing datasets; raw sequencing reads, processed data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model	.txt, .xlsx, .csv, .tsv, .fa, .bam, .rds, .h5ad, .loom, .R, .ipynb, .html, .pdf, .tif, .png, .py, .pbs	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ > 5 TB ☐ NA	

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

				☐ Software ☒ Other: sequences			
Samples	Tissue samples: whole fixed animals, injected samples, sectioned samples, frozen samples. All from octopuses bred in the laboratory.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	na	na	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☒ NA	1x -80 freezer shelf 1x -20 freezer shelf 1x 4°C fridge (depending on the type of material)
Scripts	Code written for analysis pipelines	☒ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☒ Software ☐ Other:	R scripts	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Reports	Written reports concerning the project	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	□ Audiovisual □ Images □ Sound □ Numerical ☒ Textual □ Model □ Software □ Other:	.dockx	☐ < 1 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.	I will reuse data generated previously in https://www.nature.com/articles/s41467-022-35198-1 and https://doi.org/10.7554/eLife.69161
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: No ethical approval is needed since I work with octopus embryos which are not protected by the European legislation.
Will you process personal data[4]? 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:
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:

[4] 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). *RDM guidance on documentation and metadata*.	We will maintain a record of the following for every WP for 10 years (where applicable): - Experimental design and protocol (.docx file) - Abbreviations used (.docx file) - Data structure documentation (.docx file) - Steps involved in data analysis and relevant analysis scripts (R, MATLAB, Python and ImageJ scripts) - Raw data (specific file format according to data type) - Analysed data (specific file format according to data type) - Index file/read me file (.txt file) for every WP, linking the name, location (folder and subfolder on /server) and description of above-mentioned files. Physical data: Samples taken from experiments will be documented and stored for up to three years after the end of the project. Storage will be in fixative, in paraffin, at 4C or in freezers depending on the kind of sample. Immunohistological stained slides will be stored in appropriate boxes in a dry place or fridge. We deviate from the 10 years rule because after three years, quality of the physical samples cannot be guaranteed anymore.
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 experiments are unique, but the data will be standardized according to data-type across experiments to make it easier to interpret the structure. Below, we list the metadata standards applicable to this project: Metadata standards will be used for genomics (http://enews.patricbrc.org/faqs/genome-metadata-faqs/). For all other data, metadata will be created using the Dublin core (http://dublincore.org/groups/tools/). 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-premis ☒ Large Volume Storage ☐ Digital Vault ☒ Other: ManGo (KU Leuven)
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)
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 There is currently sufficient storage at KU Leuven ICTS. If no, please specify:
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*	Researchers involved in the project can control who they give access to the files on their personal OneDrive. To access the KU Leuven servers, access is provided and controlled by the group leader, Eve Seuntjens. The KU Leuven ICTS data center hosts the network storage, with a mirror available in the second ICTS center. This ensures additional back-up capacity, recovery of lost data and long term data availability. The access is controlled by KU Leuven security groups and it is password protected.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Back-up cost per Tb (KU Leuven ICTS): 295€/year Large Volume Storage: 95,14€/Tb/year Expected amount of data (5 Tb). The costs have been budgeted on the grant

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) Digital data: We will retain all data for the expected 10 year period. We expect that we will make the data publicly available on data repositories upon publication of the manuscripts. Physical data: Samples taken from experiments will be documented and stored for up to three years after the end of the project. We deviate from the 10 years rule because after three years, quality of the physical samples cannot be guaranteed anymore.
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): Digital data will be stored at the Archive (K:) server from KU Leuven ICTS. HCR probes and physical samples will be stored in the freezers from the Research Group of Developmental Neurobiology. Code scripts will be stored on Github. Notebooks will be kept in the lab for at least 10 years.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Back-up cost per Tb (KU Leuven ICTS): 295€/year Large Volume Storage: 95,14€/Tb/year Expected amount of data (5 Tb). After the project, data preservation costs will be covered by other grants.

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: Written progress reports will be stored internally. Relevant findings will be disseminated through publication in peer-reviewed international journals. In addition, data will be presented at (inter)national scientific meetings. Transcriptomics data will be made openly available via data repositories. Requests for non-deposited data will be evaluated on a case-by-case basis and may be provided upon request.
If access is restricted, please specify who will be able to access the data and under what conditions.	All team members have access as long as they are affiliated to KU Leuven. Once all files are released, anyone can use these data to generate new results, referring to the original publication and not for commercial use. Other data will be only released upon request and after an embargo period after publication. Data will be released under a CC-BY 4.0 reuse license.
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.	☐ KU Leuven RDR ☒ Other data repository (specify) ☒ Other (specify) Experimental data will be made available through a data repository such as ncbi, github, Genbank, FigShare (https://figshare.com/), Dryad (https://datadryad.org/) or https://zenodo.org/depending on the type of data. We will explore the possibilities via online repositories and will use the website www.re3data.org.
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
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?	The transfer costs depend on the data repository selected. Costs will be covered by the project funding.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PI (Eve Seuntjens), and day-to-day manager of the FWO-project; currently: Marie Van Dijck
Who will manage data storage and backup during the research project?	The PI (Eve Seuntjens), and day-to-day manager of the FWO-project; currently: Marie Van Dijck
Who will manage data preservation and sharing?	The PI (Eve Seuntjens), and day-to-day manager of the FWO-project; currently: Marie Van Dijck
Who will update and implement this DMP?	The end responsibility for updating and implementing the DMP is with the PI, Eve Seuntjens.