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	**Pieter Vanden Berghe** ORCID: 0000-0002-0009-2094
Contributor name(s) (+ ORCID) & roles	**Candice Fung** ORCID: 0000-0002-4277-3664
Project number [1] & title	Reciprocal innervation of the gastrointestinal ileo-caeco-colonic triad; implications for storage and release of symbiotic microbial content
Funder(s) GrantID [2]	**G012223N**
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	Apart from supplying the host with nutrients and expelling undigestible remnants and noxious elements that we ingest, the gastrointestinal tract also harbors a large community of bacteria, which operate in symbiosis with its host to aid digestion and provide essential factors. The importance of the microbiome has been a matter of increased interest, and shifts in its composition have been linked to intestinal malfunction and disease. It is not clear how the gut, which mostly expels material, succeeds to keep the microbiome constant and healthy. The caecum, a small pouch at the junction of small and large intestine, functions as a fermentation reservoir, but it is at the same time a safe haven for bacteria. In this part of the intestine, a large fraction of the microbiome remains protected, which is of utmost importance for healthy symbiosis in every organism. In this proposal we aim at understanding how enteric neurocircuitry is wired in the mouse caecum, how it develops, what neurons are activated by microbial metabolites and how it connects to the other regions of the gut to assure proper function, that is filling and emptying at the right time. To this end we will make use of a bespoke dual microscopy approach that allows us to assess neuronal connectivity over long distances. Furthermore, we will also use super resolution microscopy to determine the role of the primary cilium, a small cellular antenna, which is responsible for detecting changes in the gut environment.

[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.

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].

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: P110_2020; P033_2022; P018_2022 ☐ Yes, dual use; provide approval number: ☐ No Additional information:

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

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:
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

[4] See Glossary Flemish Standard Data Management Plan

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*.	For all microscopy images: recording parameters (powers, excitation and emission wavelength), dimensions, image type, bit-depth, pixel sizes and microscope settings, will be stored. Either in a metafile accompanying the data (with identical filename) or embedded in the tiff header. The experimental protocols, stimulation settings, temperatures of the physiological experiments that will be performed during this projected will be described in detail in a lab book and referred in a ReadMe text file that will accompany the recorded data. Similarly for the processed and analyzed data, all parameters used to arrive to the results are stored within the image format.
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: Images will be stored in *.ome tiff format, which enables storing a multitude of microscope and recording parameters. For the more experimental imaging paradigms for which actual changes are made to the instrumentation, the metadata will be stored in accompanying txt or csv files. 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: All original recordings and their metadata will be stored in one copy on external harddrives, which will be labeled by projectname, subproject, experiment, data and experimenter initials. Since most of the images are highly experimental and explorative, there is no need to keep a copy of the original data. However, upon first analysis the images that are judged of sufficient quality to derive conclusive data, will be analysed as saved on the researcher's computer with a backup either on a local external harddrive or via the LUNA network. Exchange of data between the partners of the project (all KU Leuven) will happen via the LUNA network, which keeps track of versions and assures safe backups. Alternatively BELNET will be used for data transfer of files that are too large for the space available on the LUNA network. Within the FWO IRI project FBI, we are investigating future overarching datastorage and datamanagement solutions based on the IRODS / Mango pipelines.

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) As from the first steps of analysis, the data will be stored on the university's central servers with automatic daily back-up procedures or - in case too large - on local hard drives with a local backup system (e.g. Genie).
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 We have over a decade of experience recording and dealing with images and large datafiles, which we store and backup in mixed modes. For analysed images and results we use the servers (and their backup) of the LUNA network, for original data or versions of analysed image files we store locally on internal and external harddrives. Each researcher has internal as well as external PC storage and manages their data in a conscious and well thought off fashion, with necessary backups while at the same time avoiding multiple copies of less successful recordings. 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*	The data files from this study will be managed, processed, and stored on LUNA PCs and network, which can only be access by KU Leuven personnel. Access to the network storage is limited to the personnel accounts that belong to the lab of the promoter of the project and can be expanded to other reserachers involved in the project. External HD are stored in closed offices.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	16 TB external/internal HDs cost ~ 400 Euro, and thus extra storage capacity for this project is not an issue. The cost for 300 Gb storage on the LUNA network is ~ 75 Euro/y, which is all well within reach of the project budget.

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)

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): All original data that were used to generate the scientific output (papers, reports) will be stored for 10 years on external HD. These HD will be stored in closed closets in the lab of the promoter or copromoter where the data have been generated. All analysis files used for publishing and reporting will be compressed and stored on Archive directories on the LUNA network. As mentioned above, HD for the original data (of which not all is valuable enough to store in locations with multiple backups). Analysed files with relevant information will be stored on KU Leuven Archive servers conform the KU Leuven RDM policy.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	A 16 TB HD costs ~ 400 Euro which is a minimal cost compared to the running costs of the microscopy equipment and the animal housing cost. Storage for 50 Euro/TB/y = 500 Euro/TB for the storage during 10 years. We estimate 3 Tb of data will be stored on these Archive servers: 1500 Euro.

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:
If access is restricted, please specify who will be able to access the data and under what conditions.	Data will be made available either upon request of the publisher of the scientific articles or upon request via E-mail of colleague researchers.
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) Upon the publisher's request, data will be made publicly available in *.ome tiff, *.csv, or other common data formats on Zenodo or other public data servers.

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?	None. Belnet can be used to transfer and give temporary access to the requested data.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Each of the individual researchers involved in the project and the supervisor.

Who will manage data storage and backup during the research project?	Each of the individual researchers involved in the project and the supervisor.
Who will manage data preservation and sharing?	The project supervisor: Pieter Vanden Berghe.
Who will update and implement this DMP?	The end responsibility for updating and implementing the DMP is with the supervisor (promotor).