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	Rob Jelier https://orcid.org/0000-0002-6395-1407
Contributor name(s) (+ ORCID) & roles	
Project number[1] & title	**Modeling of cellular movements in a complex animal system 3E221367**
Funder(s) GrantID[2]	G008423N
Affiliation(s)	▣ KU Leuven
Please provide a short project description	Understanding how cells self-organize into complex multicellular systems is a fundamental challenge in biology. Here we propose to advance the field in a project that combines advanced microscopy, tension measurements with laser ablations, mechanical modeling of cell shapes and experimental evaluation of hypotheses. As a powerful model for complex multicellular systems, we will use the embryogenesis of the roundworm C. elegans. The embryo develops from a single cell to a free moving larvae in only 12 hours, and is nearly invariant. It is also transparent, making observations on the cellular movements straightforward. We will focus on several complex movements surrounding the early to mid-stage gastrulation. The aim is to uncover how the movements arise out of a complex mix of force generation, such as cortical tensions, the generation of pulling forces of cells on neighboring cells and cellular adhesion.

---

[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.

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].

Dataset name	Description	New or Reused	Digital or Physical	*ONLY FOR DIGITAL DATA* Data Type	*ONLY FOR DIGITAL DATA* Data Format	*ONLY FOR DIGITAL DATA* Data Volume	*ONLY FOR PHYSICAL DATA* Physical Volume
Microscopy images	Timelapses of developing *C. elegans* roundworms, made using confocal microscopy.	Both: generate new data, reuse existing data.	Digital	Experimental	.tif / .czi	< 5 TB	/
3D meshes	The result of segmenting the microscopy images. These are 3D triangle meshes representing the cell shapes.	Both: generate new data, reuse existing data.	Digital	Compiled/aggregated data	.obj	<100GB	/
Source code	Python code.	Generate new data	Digital	Software	. py	<1GB	/
Samples/strains	*C. elegans* models and samples	New and used	Physical	Experimental			Deposited in strain repository (CGC) and long-term storage in -80C freezers at CMPG

---

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

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.	Previous microscopy images and 3D meshes can be found at: https://zenodo.org/record/5108416
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:

---

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

Will you process personal data[5]? 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:
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:

---

[5] See Glossary Flemish Standard Data Management Plan

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).	Microscopy images: README.txt with image resolution, units of measurement, strain information, microscope settings. 3D meshes: README.txt with origin, cell type, units of measurement, etc. Derived data: CSV headers and optimally a README.txt for additional information. Source code: documentation (Python docstrings) included in the files. Extra guides and demos where necessary. Experimental data, strains and samples. Used and generated *C. elegans* strains are annotated with complete genotypic information. Plasmids are sequenced and annotated. Samples are primarily described in the Lab Notebooks, which are stored indefinitely, and secondarily in publications.
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: We are committed to sharing all data used in publications in a manner that stimulates further use of the data. The data are adequately described in the publications, and further in specific documentation accompanying the data as described above.

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	All source code is regularly committed to a git repository (Bitbucket). Other data are kept both locally and on shared drives of KU Leuven, as well as the professional pCloud online service for data shared with collaborators. Materials are stored in the CMPG -80C freezers, and where possible deposited in the public repositories such as AddGene and the CGC for *C. elegans* strains.
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.[6]* *REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.*	KU Leuven drives have automatic back-up facilities and are maintained by the university's IT service. All data is mirrored in a second KU Leuven datacenter. pCloud stores files on at least three server locations in a highly secure data storage area.
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: We currently have over 10TB of capacity of which more than half is still availb If no, please specify:

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?	No sensitive or confidential data is handled in this project. KU Leuven drives are protected by mandatory two-factor authentication. pCloud always requires secure connections and two-factor authentication and encryption can be used for sensitive data.
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. [6]	
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The costs are on the order of 500 euro a year and are paid from project funds, including from the current project. Storing materials in the -80C freezers is covered by the CMPG, which is funded by bench fee contributions from project funds.

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...).	Microscopy images, 3D meshes, and source code required to replicate published research will be maintained for at least 5 years. Zenodo currently guarantees at least 20 years. Samples and strains.

Where will these data be archived (stored and curated for the long-term)?	Zenodo or KU Leuven repository RDR (all data) and Bitbucket (only source code). CMPG -80C freezers.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Zenodo is free of charge. Bitbucket subscription costs about €30/ year and is covered by the research group. We use a new -80C freezer at CMPG (purchased autumn 2022) with a maintenance contract, paid for by all members with a laboratory of the center. The costs for the supervisor's lab are minimal (<100 euro a year) and are paid by project funds.

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:
If access is restricted, please specify who will be able to access the data and under what conditions.	N/A
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.	Published source code is shared at bitbucket, e.g.: https://bitbucket.org/pgmsembryogenesis/flowshape/ Microscopy images, 3D meshes and source code already available at Zenodo and KU Leuven repository RDR.

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'.*	Immediately 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."* [7]	Data: CC-BY Code: GPL-2.0
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: Standard for Zenodo.
What are the expected costs for data sharing? How will these costs be covered?	Zenodo is free of charge. Bitbucket subscription is small and covered by project funding.

---

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

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	**Rob Jelier**
Who will manage data storage and backup during the research project?	**Rob Jelier**
Who will manage data preservation and sharing?	**Rob Jelier**
Who will update and implement this DMP?	**Rob Jelier**