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	Dr. Emine Kayahan
Contributor name(s) (+ ORCID) & roles	Prof. Xavier Casadevall i Solvas (Supervisor) Prof. Ilse Smets (Supervisor)
Project number [1] & title	Intensification of light-driven water resource recovery by using droplet millifluidics for the production of seed photogranules (Insight)
Funder(s) GrantID [2]	12A6S24N
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	Over the last decade, stimulating bacteria to form spherical compact biofilms called granules instead of loose flocs has revolutionized wastewater treatment (WWT). Still, WWT technology with activated sludge granules faces some challenges such as high O_2 demand and inevitable CO_2 emissions. Recently, another type of granule, the photogranule, has been attracting attention due to its potential to achieve sustainable WWT and resource recovery. Photogranules are a mixture of heterotrophic and phototropic microorganisms. Phototrophs can autonomously produce O_2 needed while consuming CO_2 generated by heterotrophs, thereby eliminating the need for aeration of the aerobic granular technology. However, there are several problems that need to be solved to realize the full potential of this technology: 1) long start-up periods due to the non-optimized production of seed photogranules, 2) lack of fundamental studies on photogranulation and 3) poor characterization of photobioreactors (PBR) in terms of light fields. In this project I aim to address these issues by developing a droplet-based millifluidic platform to produce seed photogranules in a robust and controlled high-throughput manner. The same platform will enable me to identify the impact of environmental parameters (*e.g.,* C/N ratio and light) on the microbial composition and the resource recovery potential of the photogranules. Finally, I will validate my results in a PBR intensified in terms of light fields.

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
Microscopy images and movies	Microscopy images and movies will be taken for the development of droplet microfluidics platform, droplet generation in with PDMS chips or 3D printed pieces and the growth of photogranules inside the droplets.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☒ Other: video	.tiff .avi	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
CAD designs	The designs of the microfluidics chips will be done by Autocad. 3D printed pieces will be designed using	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☒ Model ☐ Software	.dvg .ipt .iam .par .dsm	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

	Solid Edge or Inventor.			☐ Other:			
Modelling – light transmission	Models for light transmission will be studied using MATLAB and COMSOL Ray Tracing module.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☒ Model ☐ Software ☐ Other:	.m .mat .mph	☐ < 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 not use existing data.
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:

[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: The microfluidic platform that will be developed during the project might be patentable. In this case, all the designs and the results of the platform will be kept confidential until a patent application is completed. Patent submission will be evaluated in collaboration with KU Leuven research and development (LRD) department.
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: I am using some bacteria and algae that I bought from companies like SAG or DSMZ. They have material transfer agreements (MTAs). These MTAs are not binding for publication of results as long as the company name is given as the source of the microorganisms. However, MTAs are forbidding the distribution of the microorganisms to other institutions and persons. The microfluidic platform does not need to be patented based on the microorganisms that are utilized in it. However, in case of doubt, we will check this both with the KU Leuven research and development (LRD) department and the companies that are providing the microorganisms.
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:

[4] 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). *RDM guidance on documentation and metadata.*	Protocols, description of research results, literature studies, manuscripts will be kept at as word and/or .pdf files on the KU drives. Protocols are kept on the website *"eLab Journal"*. At first, the data will be stored in based on the data type and work-packages. Later this data will be organized based on papers or patents arising from the project. All the codes and scripts will be kept in a similar manner. An example of subfolders foreseen is as follows: 1. Data –> Experimental data –> Microscopy images –> WP1 -> files 2. Data –> Codes –> MATLAB codes –> WP1-> files 3. Data –> Codes –> COMSOL codes–> WP3 -> files 4. Data –> CAD designs –> Photomask designs –> WP1 -> files See the section below for the naming conventions of the files.

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: To aid interoperability community best practices and common terminology used in the field will be used where possible. No community-agreed metadata standards exist for this discipline. Therefore, data will be appropriately documented to allow for data validation and reuse (via README files that provides information on date that the data generated, methodology used, analysis methods, variable definitions, units of measurements). To manage the data, appropriate folder structures and **file naming conventions** as below will be used: - Date or date range of experiment: YYYYMMDD - Version number of file (v1 or V001) if generated on the same day - Descriptive file name (that is still reasonably short) - Initials of the person who last modified the file if data is generated and/or modified by several people. Otherwise this information will be made available in the README.txt. - Avoid special characters (~!@#$%^&*). - Avoid spaces in file names. Use underscores (Input_File), dashes (Input-file) or camel case (InputFile).

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:
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 The total stored data is being checked by our lab manager and we discuss if we need more data during the group meetings a few times a year. In case more data storage is needed, we request it from KU Leuven. 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 network drive for the FWO-SB project folder and the large volume storage folder are secured by the ICTS service of KU Leuven with a mirror copy. Only other lab members, will have access to the shared folder. Unauthorized persons do not have access to this system.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	During the project, the data will be stored on OneDrive. KU Leuven provides 2TB storage space for free on OneDrive. In case more data storage is needed, shared network drive and large volume data storage will be considered as options. In this situation, the costs will be covered by the project.

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):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Data will be stored at Large Volume Storage of KU Leuven. The total volume of expected data is around 1 TB. Data storage costs 104,42 Euro /TB/year. For storing data for 10 years, 1044,2 Euro will be needed. The costs will be charged to the project. KU Leuven allows buying blocks of 5 TB for this storage. However, Biomimetics group has already been using Large Volume Storage. Therefore, the rest of the block 5 TB data storage space will be paid by other projects.

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.	In case of an ongoing patent application, the access to data will be restricted.
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: The sharing of the data might be delayed depending on a possible patent application. This will be discussed with KU Leuven LRD Department.
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)
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) In case of a patent application, sharing the data with open access licences might be delayed.
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?	No expected costs for data sharing are foreseen.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The postdoctoral fellow (Emine Kayahan)
Who will manage data storage and backup during the research project?	The postdoctoral fellow
Who will manage data preservation and sharing?	The postdoctoral fellow
Who will update and implement this DMP?	The postdoctoral fellow