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	**Victor Chuman – https://orcid.org/0000-0001-9719-6879**
Contributor name(s) (+ ORCID) & roles	
Project number [1] & title	**1SHC724N – High-resolution lens-free fluorescence imaging based on wavefront shaping**
Funder(s) GrantID [2]	**SB Fellowship 1SHC724N**
Affiliation(s)	☑KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☑ Other: imec ROR identifier KU Leuven: 05f950310
Please provide a short project description	The project will address the challenge of competing resolution and field-of-view in traditional microscopy by developing a High-Resolution lens-free fluorescence imaging platform based on wavefront shaping. This will be tackled by expanding the existing image formation theory to allow for better multiplexing of light on the sensor and more efficient reconstruction, and using structured illumination generated by a beam former fabricated on a photonic integrated chip (PIC). It will concentrate on the four main components of a lens-less microscope: the illumination, mask, sensor, and reconstruction algorithm. It will comprise the development of an appropriate theoretical framework, followed by the design of each microscope component, finalizing with the fabrication of a setup for experimental proof-of concept. This work builds upon imec's expertise on beam forming, structural illumination and lens-free imaging.

[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].

				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
Simulation code: 1SHC724N_SimCode_*Description_Version*	Code for the calculation and simulation of image formation and reconstruction algorithms	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☒ Software ☐ Other:	Programming language: Python - *.py, *.ipynb Julia - *.jl	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Simulation images results (Own code): 1SHC724N_SimResults_*Description_Version*	Images generated with the self-written code	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	Image format: *.png, *.tif, *.tiff, *.hdf5	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

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

				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
Simulation results (FDTD, FEM, BEM): 1SHC724N_SimResults_*Method_Description_Version*	Results from running 3rd party physics simulation suites	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	*.tif, *.tiff, *.hdf5	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
CAD designs: 1SHC724N_Design_*Description_Version* (The software is evident from the extension)	Designs required for the manufacturing of the experimental setup	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☒ Model ☐ Software ☐ Other:	*.gds, *.step, *.sldprt, *.slddrw, *.sldasm, *.pdf	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Experimental data: 1SHC724N_Experiment_*Description_Setup_Date_Trial*	Data acquired during the experimental proof-of-concept of the developed imaging approaches	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	*.tif, *.tiff, *.hdf5	☐ < 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.	
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:
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: Simulation code: Aid in designing functional imaging set-ups. CAD designs: Though proof-of-concept, they could be potentially ripened to design a viable product.

[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*.	An **excel data catalog** will be used and regularly updated. A folder and filing hierarchy will be set out and documented, keeping the closet resemblance possible to the project plan. At the end of the project, the data will be reviewed on reusability, and will be added to the data catalog excel file in the main folder. Anywhere dimensions are relevant, SI units will be used. **Simulation code:** Core calculation functionality is programmed and commented in the source code. Notebooks (e.g. Jupyter) are used to keep all information necessary to understand and reuse the code. Reference to any additional and necessary data are made directly in the notebook. **Simulation images and experimental images:** Images not containing metadata (*.png) will be named with a unique identifier and a readme file/table will be added to the folder. Images containing metadata (*.tif, *tiff, *.hdf5) will contain the necessary information in the metadata. **Simulation models and results, CAD designs:** Unique identifiers will be used. The software annotation capabilities of the corresponding suite will be used. Were not possible, an additional readme file/table will be used. **For experiments:** A digital **lab notebook** will be kept in OneNote, where the specific experiment(s) are described with more detailed parameters (time, consignee, protocol, samples names, conditions, ...). **Metadata** will be provided following (where and if applicable) an adapted Dublin Core Metadata standard with the following metadata elements: Data ID, title, subject, name, creator, related project ID, data abstract, keywords, type, size, format, source, related publications or patents, compliance, security and confidentiality aspects, rights. **For experiments**, the following metadata elements will be added where and if applicable: Experimental set-up, study design, sampling methodology, variable-level detail, reference to specific model.

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: Where relevant, an adapted Dublin Core Metadata standard will be used. The fields are already described above.

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: imec sharepoint, imec OneDrive and imec data server units

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) During the project data will benefit from an automatic back-up: - This is managed by Microsoft: The imec sharepoint is hosted on the Microsoft Cloud, which is a high available environment. Information on this environment is never automatically deleted, removal of specific information should be managed by the business. - If information is removed from this environment, it is moved to a "Recycle Bin". From this Recycle Bin it can still be restored by the user for a period of 93 days (or an admin in case the Recycle Bin was emptied). - After these 93 days, items are deleted, and Microsoft will keep a backup for 14 additional days. During this period restoration can still be requested through a Microsoft ticket. - After these 14 additional days, the data is permanently removed. Other data on premises (Isilon) have a automatic back-up system. Data is backed up via snapshot technology, where all incremental changes in respect of the previous version are kept online.
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, at imec data servers, which are expanded at any point it is necessary. ☐ No 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*	Data on SharePoint will benefit from the following security services: - User-authentication, multifactor authentication can be activated. • All logins to imec environments are protected by a strong password (minimal 14 characters) combined with multi factor authentication. - Versioning system • Versioning is enabled by default on our SharePoint/Teams' environment. This can be modified by workspace owner and is under full responsibility by the business user. - System-encryption • This is managed by Microsoft as outlined in this article: Cloud data security measures in SharePoint & OneDrive - SharePoint in Microsoft 365 | Microsoft Docs • Making sure that data is shared with the right people is the responsibility by the business user. • Confidential data will be user encrypted in addition to this: ▪ When creating documents (.docx, .xlsx, .pptx), the user is forced to assign an information classification label to the document (Public / Restricted / Confidential / Strictly Confidential) based on the sensitivity of the information. ▪ While this label provides a (visual) marker on the sensitivity level of the information, it will not encrypt the document by default.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Storage is available at imec at no added cost to the project. In general, can be calculated with 50€ per month for 1 TB of data. However, this will not be being charged internally and is included in overhead expenses.

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): imec sharepoint, OneDrive and data server units. After the standard retention period of 5 years, data will be subjected to evaluation. This will include weighing the potential value versus the costs of keeping it available. Decisions will be made by the data owners, in close collaboration with ICT service responsibles for archiving.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	In general can be calculated with 50€ per month for 1terrabyte of data. However, this will not be being charged internally and is included in imec overhead expenses.

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: Some of the data can be provided to a third party under reasonable request (e.g. for publications or other collaboration possibilities).
If access is restricted, please specify who will be able to access the data and under what conditions.	Internal imec employees related to the project can have access. Any external employee could get partial access under a reasonable request (e.g. for publications or other collaboration possibilities). We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. It is expected to obtain data that can potentially be valorised in all workpackages. The data will be kept confidential until 1) a decision will have been made on patenting, and 2) a patent file or publication has been submitted.
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: Part of the data might lead into potential industrial applications. The authors involved in the generation of the data (e.g. imec, KUL) as holder of the intellectual property, might restrict access to third party.

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) Zenodo.org will be the preferred standard repository for open data, because this platform, part of the OpenAIRE collaboration, provides the minimum and recommended terms required by DataCite's Metadata Schema, including the assignment of an identifier. These services are furthermore free of charge and will be for the foreseeable future. Software code will be placed in GitHub.com in an open accessible manner to the public. Physical data and samples will only be stored until the intended analyses are finished and all information is gathered.
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify) Upon publication, if it does not conflict with ongoing intellectual property protection procedures.
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?	These costs will be covered in project overhead, except in case of circumstances that can be considered out of the ordinary.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	**Victor Chuman, Niels Verellen**
Who will manage data storage and backup during the research project?	**Victor Chuman, Niels Verellen**
Who will manage data preservation and sharing?	**Victor Chuman, Niels Verellen**
Who will update and implement this DMP?	**Victor Chuman, Niels Verellen**