

Expanding the chemistry toolbox of next generation Expansion Microscopy: Ex2

FWO DMP (Flemish Standard DMP)

1. 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.

				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
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> Generate new data Reuse existing data 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> Digital Physical 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> Observational Experimental Compiled/aggregated data Simulation data Software Other NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> .por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ... NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <100MB <1GB <100GB <1TB <5TB <10TB <50TB >50TB NA 	
Microscopy images	Raw microscopy datafiles	Generate new data	Digital	<ul style="list-style-type: none"> - Observational - Experimental - Compiled/aggregated data - Simulation data 	.por, .xml, .tab, .csv, .pdf, .rtf, .dwg, .gml, ...	<1TB	
Experimental protocols	Organic synthetic protocols and labelling protocols	Generate new data	Physical				Lab notebooks
Data processing	Figures, datasets, presentations, research article drafts	Generate new data	Digital	<ul style="list-style-type: none"> - Observational - Experimental - Compiled/aggregated data 	.pdf, .docx, .xlsx, .pptx, .png, .jpeg, .cdxml	<1TB	
Synthetic experimental data	Raw datafiles containing compound characterisation (NMR, LC-MS, ...)	Generate new data	Digital	<ul style="list-style-type: none"> - Observational - Experimental - Compiled/aggregated data 	NMR data, LC-MS data	<1TB	

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:

No

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

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

The methods and materials developed during this project have potential for valorization, which will be a point of discussion during the project. Patents and tech transfers will be submitted, transferred and valorized with guidance of the intellectual property unit of KU Leuven Research and Development (LRD).

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 in the comment section to what data they relate and what restrictions are in place.

- No

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

2. 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).

Details to reproduce organic reaction procedures are described in a personal lab notebook (date, unambiguous identifier, reaction description, molar quantities, reaction protocols, work-up protocols, purification protocols, further suggestions and unambiguous identifiers of characterization data) as well as on digital sources (word documents). ReadME-files or recording methods for characterization (NMR and LC-MS) are stored in parallel with generated data.

Biological labelling procedures (date, sample description, used labelling compounds, concentrations, labelling ratio's, protocols, experimental set-ups, further suggestions and short conclusions) described in a personal lab notebook, as well as on digital sources (word documents). ReadME-files or recording methods for produced microscopy images are stored in parallel with generated data.

For presentations and publications, the procedures are described in sufficient detail to reproduce the data for experienced users.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

No metadata standard will be used. The measurement setup will automatically generate metadata which is embedded within the characterization (NMR, LC-MS) and microscopy data sets. The data will be stored and named in consistent manner with unambiguous identifiers. Descriptions of experiments and materials are kept in physical lab notes for each identifier. The data will be processed and summarized in comprehensive word/excel files with reference to the identifiers. The resulting summaries will culminate in presentations and publications/patents.

3. Data storage & back-up during the research project

Where will the data be stored?

Physical lab notebooks of organic synthetic procedures and labeling protocols.

The digital data is stored on personal portable hard drives, some equipment used to take such data as well as the KU Leuven cloud service 'One Drive', which offers secure storage up to 1TB for each user.

The accepted version of the final manuscripts (+ accessory datasets and supporting information) are submitted in open access journals and in the KU Leuven library's depository. In addition, relevant data will be transferred to the self-mirrored long-term storage capacity by the KUL ICTS for long-term storage till at least 5 years after the project

How will the data be backed up?

Digital data will be stored in parallel on the cloud (one drive) and on the personal computer/personal portable hard drive of the applicant, guaranteeing backup of the data.

The physical data of the lab notebooks is often transferred to digital data, following the storage method mentioned above.

**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

Storage and backup capacity is already present. In any case it can be extended further by purchasing additional portable hard drives or cloud data volumes.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Access to the data is restricted by online build-in safe logins and by physical restriction of access to the portable hard drives (through the security system installed in the building of the affiliation).

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

Expected additional costs remain limited, as data storage is ensured.

Additional costs may include the purchasing of further portable hard drives (2TB for 100-200 euro), which will then be covered by the personal bench fee. Cloud-based data storage using one drive software volume will be shared with the host research group.

4. 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...).

Only a selection of relevant data will be preserved. Selection of data destined for long-term storage will be assessed by the applicant and the promoter on a yearly basis.

Naturally, all data and metadata used/linked to published articles, will be preserved in the university library depository.

Physical lab samples have a shelf life and take up valuable space making them difficult to store over long periods of time.

Where will these data be archived (stored and curated for the long-term)?

The selected data will be stored on the university's central servers for at least 10 years, confirm the KU Leuven RDM policy. In addition, the selected data will be preserved on minimum 3 different portable hard drives which are kept locked at different locations under the supervision of the promotor.

The physical lab notebooks will also be stored ...

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

Costs for the retention of data on portable hard drives is already made. The long-term data storage of ICTS costs approximately 150 euros per year per 2TB. This storage has been set up for the research group, the shared cost will depend on the actual data generated from this project. A volume of 5 TB is anticipated, that will result in a cost of approximately 300 Euros to be covered from the bench fee of this grant.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, ...)

If access is restricted, please specify who will be able to access the data and under what conditions.

The data will continue to be stored in a protected environment, similar as to before the end of the project. However, requests to access the data can be made via email to the promotor or PhD student, who will decide upon this request after consultation with the promotor and co-authors of articles and after the possible receiver has signed a data sharing agreement.

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 in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Data will be made available through publications and the accompanying supporting information. Additional characterization, microscopy or protocol data can be made available upon request to myself or PI via email.

When will the data be made available?

Upon request to myself or PI via email as well as upon publication of research results.

Which data usage licenses are you going to provide? If none, please explain why.

Creative Commons Attribution-No-Derivs (CC-BY-ND)

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

How data will be shared, will depend upon the type of the requested data. Data will always be shared in a protected environment; after the receiver signed a data sharing agreement.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Thibo Iven

Who will manage data storage and backup during the research project?

Thibo Iven

Who will manage data preservation and sharing?

Thibo Iven and Johan Hofkens

Who will update and implement this DMP?

Thibo Iven