Smart integration of membrane technology and polymer chemistry as scalable access to novel groundbreaking macromolecular materials

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.

The project will generate experimental datasets originating from the synthesis of sequence-defined macromolecules at UGent and KUL2, as well as the synthesis and characterization of membranes for nanofiltration (KUL2).
Laboratory notebooks, reports, publications and data extracted from literature survey will be performed by all partners (UGent, KUL1 & 2).

The main techniques to characterize the sequence-defined macromolecules, will be performed by UGent and KUL2, and these consist of: Nuclear magnetic resonance spectroscopy (NMR), Fourier transform infrared spectroscopy (FTIR), Size exclusion chromatography (SEC), MALDI-TOF, Liquid-chromatography mass spectrometry (LC-MS).

Additionally, UV-Vis will be utilized and performed by all partners.

				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
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .cvs,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Lab notes	Description of the practical execution of experiments	New	Digital and if Physical, they will be digitalized as materials and method section	Observational and experimental	.docx	<100 GB	5-10 note books
NMR	NMR spectra	New	Digital	Experimental	NA	< 100 GB	
UV-vis	UV-vis spectra	New	Digital	Experimental	.xlsx	< 100 GB	
Filtration results	Results form the filtration experiments	New	Digital	Experimental	.xlsx	< 100 GB	
Literature	Background literature	Existing	Digital	Other	.pdf	< 100 GB	
IR	infrared spectroscopy data	New	Digital	Experimental	.xlsx/.dpt/.0	< 100 GB	
Maldi-ToF	Mass spectra	New	Digital	Experimental	.xlsx / .text	< 100 GB	
SEC	Size exclusion chromatogram	New	Digital	Experimental	.xlsx / .text	< 100 GB	
LC-MS	Liquid chromatography coupled to mass analysis	New	Digital	Experimental	.xlsx / .text	< 100 GB	

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:

Data published in the literature will be reused but no specific DOI can be given at this stage. These literature will be clearly identified in the laboratory notebooks and in the reference section of our publications.

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

Before every publication, the potential towards IP creation will be assessed. If there is IP potential, the relevance will be communicated with the relevant tech transfer offices. Therefore, all relevant data will be kept restricted until assessment towards filing a patent. Once these granted, this data can be made public.

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

There is no third party agreement restricting data dissemination or exploitation, and no specific agreement between the two institutions.

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).

The data files will be named using a standardized naming system, including date of the experiment, name of the researcher, sample code,... The used codes will correspond to the codes used in the lab notebooks.
An index or table of content file will be provided with the explanation of each code and a short description of each related project. In this index, also a link will be embedded to the data file location.
Instruments settings and SOPs procedures are made available to the 3 partners on the project's SharePoint.

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

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

Where will the data be stored?

All collaborative data generated will be made available to all the involved researcher's via the project's SharePoint (Accessible to all involved researcher via MS Teams).
Specific data generated by KULeuven partners (i.e.; related to the conjugated sequence-defined oligomers and nanofiltration) will be stored on a OneDrive folder shared among KU Leuven researchers. At UGent, data, linked to the synthesis of non-conjugated sequence-defined macromolecules, are stored on the POLCHEM server (belonging to the research group and not dependent on a researcher's contract, therefore having no time restriction), as well as on the 2Tb personal OneDrive folders provided by the University to each researcher.

Each partner is therefore responsible for the storage of the data generated by their own research group.

How will the data be backed up?

Non-digital data (lab notes) will be backed-up by digitization of the enclosed data. Digital data will be stored on both the central network drives and on the KULeuven (Archive) and UGent (POLCHEM server). The servers are not deleted after departure of the researchers as they belong to the research groups

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

The KUL and UGent One Drive has 2 TB storage capacity. This should be sufficient to store all data derived from this project.

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

The lab notes and work laptop are stored in a locked cupboard in the researcher's office. The office is located in a badge-restricted area of the building, and is locked if no one is inside.
The laptop is secured with a password and access to double authentication is required to access the KULeuven and UGent One Drive system from other devices.
OneDrive ensures that each file is encrypted at rest with a unique AES256key. These unique keys are encrypted with a set of master keys that are stored in Azure Key Vault of Microsoft.
No very sensitive data will be generated, therefore no extra security (encryption of the computer) is foreseen at this stage.

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

OneDrive and Teams are provided by KU Leuven and UGent. For KU Leuven, the internal storage costs are estimated to be 50 euro per TB for backup on the internal severs. These costs will be covered by the involved research groups.

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...).

All data will be preserved for 10 years after the end of the project according to KU Leuven and UGent RDM policy.

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

The data will be archived on the central KULeuven (Archive) and UGent (POLCHEM) servers.
These are backed up monthly on a secure offline server.

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

The costs for long term data storage are 50 euro per TB per year.
It's expected that the maximum amount of archived data will be less than 1 TB.
The maximum costs for data preservation will thus be 500 euro, which will be covered by general lab funding.

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, …)

Most of the research project involves the synthesis and optimization of procedures and protocols that are of utmost importance to the research groups. These data for instance could result in Standard Operating Procedures (SOPs) and synthesis protocols that are used and shared within the research groups.
Data will be made available upon publication. Unpublished data can be requested at one of the PIs (Ivo Vankelecom, Filip Du Prez or Guy Koeckelberghs).

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

All researchers and PI will have access at all time to the data. Externals can get access to the data upon approval of the PI.

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

Only data with potential IP protection will be restricted to the consortium members and will not be published before filing a patent.

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

Publications will be found open access on FRIS, KU Leuven RDR and Biblio at UGent. Non published datasets are available upon request to external and will be available for project members via the MS Teams SharePoint.

When will the data be made available?

Upon publication of the results

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

Data will be first restricted with an embargo (in case a patent is filled in) and afterwards data will become open under a Creative Commons license.

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?

FRIS is free. KU Leuven RDR free for 50 GB. UGent Biblio is free. Other institutional servers for non-published data are free of charge.

6. Responsibilities

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

Sareh Rezaei Hosseinabadi, Sutapa Roy Swarna, Zahra Bozorgmehrc, Irene De Franceschi, Wout Milis

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

Sareh Rezaei Hosseinabadi, Sutapa Roy Swarna, Zahra Bozorgmehrc, Irene De Franceschi, Wout Milis and Nezha Badi.

Who will manage data preservation and sharing?

Ivo Vankelecom, Filip Du Prez, Guy Koeckelberghs, Nezha Badi and Laurens Rutgeerts

Who will update and implement this DMP?

Laurens Rutgeerts and Nezha Badi

Smart integration of membrane technology and polymer chemistry as scalable access to novel groundbreaking macromolecular materials
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

The project will generate experimental data from compound synthesis, research protocols and physico-chemical analyses.
The first dataset types that will be generated are originating from the analysis of small chemical compounds, polymers and membranes.
A second type of data that will be generated during the project will be numerical and multimedia files (e.g. images and videos of experiments).
A third relevant dataset will include all the research protocols, lab notes, PowerPoint presentations, reports and publications.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

Data will be stored 10 years after the end of the project.
Data with long-term value will be mostly published as research articles in journals that often support Open Access. Should articles not be published open access, these will be available through the open access repository platforms from UGent and KUL.
When results are published, the intermediate data and the recorded workflows that enable accurate and faster manuscript writing, will be published alongside with the original research data. Other data will be stored on specific SharePoint platforms.
Physical data carriers (e.g. laboratory notebooks) are archived on site. At UGent, all laboratory notebooks are digitalized and are stored on the internal UGent server (Polchem).

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

Not applicable

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

No

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

None

Smart integration of membrane technology and polymer chemistry as scalable access to novel groundbreaking macromolecular materials
DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

Smart integration of membrane technology and polymer chemistry as scalable access to novel groundbreaking macromolecular materials

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable