

## Multiple diagnostics platforms for in situ crystallization studies connected by differential impedance measurements (DIM-C)

### DPIA

---

#### DPIA

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

- Not applicable

## Multiple diagnostics platforms for in situ crystallization studies connected by differential impedance measurements (DIM-C)

### GDPR

---

#### GDPR

Have you registered personal data processing activities for this project?

- No

# Multiple diagnostics platforms for in situ crystallization studies connected by differential impedance measurements (DIM-C)

## 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)**

Textual, image and video data will be stored in open accessible file formats. Textual data will be stored as doc, docx, .pdf formats. Image data will be stored as .jpeg and .tiff for bitmap images and .svg/.pdf for vector images. Videos will be stored as .mp4. No audio data will be generated during the project.

The automation and data acquisition software for custom setups will be developed using Python, a widely known programming language with OSI-approved open-source license, and saved in corresponding .py files. Code documentation will be integrated directly within the program. More elaborate programs and software tools will be documented in manuals, either saved as .docx or .pdf.

Data provenance will be provided in the README.txt file, which will be deposited together with datasets. Characterization techniques expected to be used and the formats of corresponding numeric data are shown in the table below. For exported data, open-access data formats will be used, as far as possible.

Characterisation technique data formats:

Impedance data .txt, .csv

X-ray diffraction/scattering STOE: .raw / ESRF: .edf

Solid / liquid state NMR Bruker .ser, .fid

**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)**

1. Responsible person: Dr. Eric Breynaert
2. Storage capacity/repository
  - All data generated during the project will be stored centrally on storage facilities of the KU Leuven which provides daily automatic back up and is maintained by the IT staff of the university. KU Leuven provides 2 TB of storage for free for every employee.
  - Upon publication and after the research, datasets will be deposited in Harvard Dataverse (1TB per researcher). The Harvard Dataverse is freely accessible. professional platform carefully developed to make data accessible according to the FAIR principles.All features needed to make the project compliant with the FAIR principles are readily available on this platform.

**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)**

There are no reasons to deviate from the principle of preservation of data.

**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)**

-



Multiple diagnostics platforms for in situ crystallization studies connected by differential impedance measurements (DIM-C)
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
		Please choose from the following options: <ul style="list-style-type: none"> <li>Generate new data</li> <li>Reuse existing data</li> </ul>	Please choose from the following options: <ul style="list-style-type: none"> <li>Digital</li> <li>Physical</li> </ul>	Please choose from the following options: <ul style="list-style-type: none"> <li>Observational</li> <li>Experimental</li> <li>Compiled/aggregated data</li> <li>Simulation data</li> <li>Software</li> <li>Other</li> <li>NA</li> </ul>	Please choose from the following options: <ul style="list-style-type: none"> <li>.por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ...</li> <li>NA</li> </ul>	Please choose from the following options: <ul style="list-style-type: none"> <li>&lt;100MB</li> <li>&lt;1GB</li> <li>&lt;100GB</li> <li>&lt;1TB</li> <li>&lt;5TB</li> <li>&lt;10TB</li> <li>&lt;50TB</li> <li>&gt;50TB</li> <li>NA</li> </ul>	
IS	Impedance spectroscopy data	Generate new data	Digital	Experimental	.csv, .txt	<100GB	
XRD	X-ray diffraction data	Generate new data	Digital	Experimental	.raw, .edf	<5TB	
NMR	NMR data	Generate new data	Digital	Experimental	.ser, .fid	<1TB	
COMSIM	COMSOL simulation data	Generate new data	Digital	Experimental	.mph	<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:

It is not planned to reuse 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)? 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.

- No

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).

I commit to implementing the FAIR principles for managing DIM-C data, making it Findable (F), Accessible (A), Interoperable (I), and Reusable (R) at the earliest possible stage. With respect to the Findable & Accessible, the results produced by DIM-C will be published in open-access, peer-reviewed journals (see Section 2.2.1 of proposal). Data will be deposited in Research Data Repository (RDR), a recently established local repository hosted by KU Leuven, and linked or uploaded to Zenodo and the Harvard Dataverse. Datasets will receive Digital Object Identifiers (DOIs), enabling citation in papers and facilitating links to online science communication platforms such as ResearchGate. Deposited datasets will include COMSOL models and CAD drawings of the simulated setups, experimental data, as well as metadata thoroughly describing the experiments & samples, their history, and measurement conditions. To ensure the published data is Interoperable, it will be provided in freely reusable formats such as txt, csv, jpeg, ascii dat, jcam-dx, or cif, whenever possible. Reusability will be guaranteed by linking published datasets to the original papers, which will also be available in Open Access (OA). KU Leuven supports Green OA via the institutional repository Lirias<sup>28</sup>. It also promotes the use of the Open Research Europe publishing platform and has a read and publish agreement with the Royal Society of Chemistry covering article processing charges (APCs) in hybrid RSC journals. Unpublished work will always be accessible through my supervisor dr. E. Breynaert, with release following an embargo period. This embargo (max. 5 years) will depend on the publication timeline. Upon publication, the embargo will immediately be released and associated data will be made openly accessible. The KU Leuven OA Desk ensures copyright compliance and assists in providing OA to the papers reporting on data in the RDR.

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 data generated during the project will be stored centrally on storage facilities of the KU Leuven. KU Leuven provides 2 TB of storage for free for every employee.

#### How will the data be backed up?

KU Leuven provides daily automatic back up, maintained by the IT staff of the university.

**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

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

Accessing data online requires two-factor authentication managed by KU Leuven. Computers, where data is stored in mirrored folders, are secured with biometric login or PIN codes.

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

< 1000€, costs will be covered by KU Leuven.

### 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...).**

Yes, data will be retained for at least five years.

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

Harvard data repository

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

No costs are expected

## 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 an Open Access repository

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

Not 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 in the comment section per dataset or data type where appropriate.**

- No

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

Harvard Dataverse

**When will the data be made available?**

Unpublished work will always be accessible through my supervisor dr. E. Breynaert, with release following an embargo period. This embargo (max. 5 years) will depend on the publication timeline. Upon publication, the embargo will immediately be released and associated data will be made openly accessible. The KU Leuven OA Desk ensures copyright compliance and assists in providing OA to the papers reporting on data in the RDR.

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

CC BY

**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.**

- Yes

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

I will opt mainly for free services such as the Harvard Dataverse, so costs are expected.

## 6. Responsibilities

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



Nikolaus Doppelhammer

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

KU Leuven

**Who will manage data preservation and sharing?**

Nikolaus Doppelhammer, Eric Breynaert

**Who will update and implement this DMP?**

Nikolaus Doppelhammer, Eric Breynaert