SPARKLES - Femtosecond laser processing for micro- and nanopatterning of metal halide perovskite thin films for enhanced light management in solar cells
DPIA

DPIA

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

- Not applicable

GDPR record

Have you registered personal data processing activities for this project?

- Not applicable

SPARKLES - Femtosecond laser processing for micro- and nanopatterning of metal halide perovskite thin films for enhanced light management in solar cells

Created using DMPonline.be. Last modified 03 October 2024

2 of 12

SPARKLES - Femtosecond laser processing for micro- and nanopatterning of metal halide perovskite thin films for enhanced light management in solar cells
Grant proposal

1. Data summary

1.1 Types of data/other research outputs

Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume	Physical Volume
Samples	Laser patterned thin film	New data	Physical	-	-	-	Few hundred samples of 2x2 cm
Researcher's notebooks	Notebooks used during the experiments/results discussions	New data	Physical	-	-	-	Around 2 notebooks per year
Experimental data	Raw data from the different characterization techniques (mainly: Fluorescence, SEM, Raman spectroscopy, UV-VIS, XRD, laser recipies)	New data	Digital	Experimental	.csv .txt .oif .tiff .jpg .lif .rcp	< 500 Gb	-
Analysed data	Reports, figures, presentations, data mining	New data	Digital	Compiled data	.pdf .txt .pptx	< 10 GB	-
Electronic lab notebooks	Short description of experiments and researcher observations	New data	Digital	Compiled data	.txt .pdf	< 1 GB	-
Meeting summaries	Summary of meetings with supervisors	New data	Digital	Compiled data	.pdf	< 100 MB	-

2. FAIR principles

2.1 Findability of data/research outputs

The digital data will be stored in the Research Data Repository (RDR) of KU Leuven after processing. This institutional platform enables the detailed description of research datasets and provides a Digital Object Identifier (DOI) for the uploaded files, ensuring the data's findability.

2.2 Accessibility of data/research outputs

The data will be accesible througt RDR repository under the open access licence (Creative Common Attribution International Public Licence (CC-BY)). The data will be also stored in local hard drive in the Roeffaers group, and will be accesible to the researchers continuing this research line.
The peer-reviewed publication will be deposited in LIRIAS (KU Leuven repository) as soon as possible.

2.3 Interoperability of data/research outputs

Standard terminology and notation will be used during all the project. To avoid any confusion, all reports generated in the project will clarify any abbreviations and units, mantaining the physical meaning of the parameters.

2.4 Reusability of data/research outputs

The main results from the project will be published under the Creative Commons CC-BY license as soon as possible, taking the publication delays into account. These documents will include the most relevant experimental details ensuring the reproducibility of the results.

All the dataset generated during the research will be labeled descriptively, including fabrication details of the sample (e.g. date_material_synthesis conditions_processing parameters), in such a way that all the data can be re-analyzed in the future if needed and the samples can be reproduced.

An electronic lab notebook will be created for each experiment, containing a brief description of the sample (including synthesis and processing parameters), the goal of the experiment and the main observations of the researcher.

The fabrication protocol developed during the project will be written and continuously updated. These reports will be stored internally as Standard Operating Procedure (SOPs) for future researchers and will be shared under request for external researchers.

3. Resources and responsibilities

3.1 Curation and storage/preservation costs

All the electronics data generated during the project will be stored on local computers just as OneDrive folder. Regarding the physical data, the samples will be stored in the proper place considering the special conditions required for each type of sample (glovebox, desiccator, regular box, etc.). The researcher's notebook will be accesible in the office of the researchers.

The digital data backups will be saved periodically in the KU Leuven network K-drive of the Roeffaers group ((\\ICTS-S-DFS5.luna.kuleuven.be\archive\SET-cMACS-ResearchData-D0758). All the data will be preserved during and 5 years after the project.

The amount of digital storage capacity expected in this project is easily addressed with regular capacity of local computers and K-drive space. No extra storage cost is expected.

3.2 Person/team responsible for data management and quality assurance

Dr. Rocío Ariza García is the responsible person for managing the research data and outputs.

For long term storage, the responsible person is Prof. Maarten Roeffaers.

SPARKLES - Femtosecond laser processing for micro- and nanopatterning of metal halide perovskite thin films for enhanced light management in solar cells
Full DMP

Version information

Action number

101149132

Action acronym

SPARKLES

Action title

Femtosecond laser processing for micro- and nanopatterning of metal halide perovskite thin films for enhanced light management in solar cells

DMP version number

1.0

Date

24/09/2024 - Initial DMP

1. Data summary

1.1 Will you re-use any existing data and what will you re-use it for?

No existing data will be re-use in this project.

1.2 What types and formats of data and other research outputs will the project generate or re-use?

Digital data type will be mostly numerical data. The envisaged data object are spreadsheets, databases, images, maps. The mode of data collection are through experiments.
The collected data will be mainly born-digital, but can also be complemented by non-digital notes.
The data life cycle will go through the following stages: raw data, processing and data output for figures in manuscripts. The collected data will be saved in the most simple, open file format, such as .txt or .csv file formats, most commonly used by researchers. The original formats (licensed software) can be shared upon request.
The type of data organized by work packages are:
WP1: Training through research and project management

- Word/pdf reports (e.g. DMP and Career development plan)

WP2: MHP synthesis and optimization

- Reports (.txt / .pdf)

WP3: Advanced laser patterning of MHP thin films and relevant substrates

- Laser setup files (.rcp). As this format is only compatible with the laser software, the laser processing parameters and the

sample "map" will be created in powerpoint (.ppt). It should be emphasized that these conditions are entirely dependent on each experimental system.

WP4: Materials and device characterization and optimization

- Characterization techniques: SEM (.tif), fluorescence (.csv), FLIM (.tif and .csv), UV-VIS (.csv), PL(.csv), Raman spectroscopy (.csv).

WP5: Design and optimization of laser-patterned MHP solar cells

- Reports regarding the device architecture (.txt / .pdf)
- Performance testing (.cvs)

Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume	Physical Volume
Samples	Laser patterned thin film	New data	Physical	-	-	-	Few hundred samples of 2x2 cm
Researcher's notebooks	Notebooks used during the experiments/results discussions	New data	Physical	-	-	-	Around 2 notebooks per year
Experimental data	Raw data from the different characterization techniques (mainly: Fluorescence, SEM, Raman spectroscopy, UV-VIS, XRD, laser recipies)	New data	Digital	Experimental	.csv .txt .oif .tiff .jpg .lif .rcp	< 500 Gb	-
Analysed data	Reports, figures, presentations, data mining	New data	Digital	Compiled data	.pdf .txt .pptx	< 10 GB	-
Electronic lab notebooks	Short description of experiments and researcher observations	New data	Digital	Compiled data	.txt .pdf	< 1 GB	-
Meeting summaries	Summary of meetings with supervisors	New data	Digital	Compiled data	.pdf	< 100 MB	-

1.3 What is the purpose of the data generation or re-use and its relation to the objectives of the project?

The main objective of the project is patterning metal halide perovskite thin films with fs-laser. It is critical a well-understanding of the influence of the laser in the thin film. For this reason, different fs-laser processing conditions will be explored and a significant amount of data related with the characterization will be generated.
The data obtained during this project form the basis for understanding future results and are crucial for the continuation of the research line.

1.4 What is the expected size of the data that you intend to generate or re-use?

A detailed estimation is given in 1.2.
Total storage of digital data is estimated to be less than 500 Gb including all the raw data (some microscopy techniques create file with large sizes) and reports.

1.5 What is the origin/provenance of the data, either generated or re-used?

All the data will be generated from scratch. The expected techniques that will be employed in the research are: SEM, Fluorescence microscopies (+FLIM imaging), Raman spectroscopy, UV-VIS, PL, Synchrotron measurements, I-V curves, degradation tracking (optical microscopy).

All the techniques can generate files in (.csv) which can be open/processed in any opensource software. The data analysis will be done using Origin (the research group has several licences for using) but the results will be saved in general .csv format for sharing.

1.6 To whom might your data be useful ('data utility'), outside your project?

The generated data could be of interest to researchers focused on light-management strategies and perovskite patterning for different applications. As the main purpose of this project is to develop new and advanced light management strategies for solar cell application using metal halide perovskite, the results may also attract interest from specialized industry.

2.1 FAIR data: Making data findable, including provisions for metadata

2.1.1 Will data and other research outputs be identified by a persistent identifier?

- Yes: describe below

Relevant data will be deposited in RDR, the KU Leuven reposity, that provide a DOI for the identification of the data.
All the collaborators in the project will be clearly identified by their ORCID number.

2.1.2 Will rich metadata be provided to allow discovery?
What metadata will be created?
What disciplinary or general standards will be followed?
In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

There is no metadata standard in the reseach field. Nevertheless, a specific protocol has been developed for ensuring all the data are accessible and easy to understand and reuse. This protocol include a specific template for labeling, synthesis and laser parameter info in the electronic labnote for an easy findability.

2.1.3 Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

- Yes: describe below

The data will be organized in characterization techniques, making easier the creation of keywords for optimizing the findability and re-use of the data.
The keywords will be the name of the technique.

2.1.4 Will metadata be offered in such a way that it can be harvested and indexed?

- Yes: describe below

RDR platform allows to include the most relevant metadata and this platform provides a DOI number with all the data and metadata.

2.2 FAIR data: Making data accessible

2.2.1 Will the data and other research outputs be deposited in a trusted repository?

- Yes: describe below

The digital data will be stored in the Research Data Repository (RDR) of KU Leuven after processing. This institutional platform enables the detailed description of research datasets and provides a Digital Object Identifier (DOI) for the uploaded files, ensuring the data's findability.

2.2.2 Have you explored appropriate arrangements with the identified repository where your data and other research outputs will be deposited?

- Yes

RDR repository is a trusted place for the data deposition and KU Leuven guarantee the suitability of the data.

2.2.3 Does the repository ensure that the data and other research outputs are assigned an identifier? Will the repository resolve the identifier to a digital object?

RDR will provide a DOI number for all the data deposited.

2.2.4 Will all data and other research outputs be made openly available?

- Yes

The publications and the data that support the results will be publish in an open access repository as soon as possible. This data will be mainly the synthesis and processing methods.
If some of the results are potentially patented, the access to those data will be considered and evaluated with the KU Leuven tech transfer office (LRD). KU Leuven RDR offers restricted access to data in these cases.

2.2.5 Is an embargo applied to give time to publish or seek protection of the intellectual property (e.g. patents)?

- Yes

2.2.6 If an embargo is applied (see question 2.2.5), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

The possibility of embargoes will be carefully analyzed with LRD. In this case, the data will be published in restricted RDR, and will be made open access as soon as the IP situation is clarified.

2.2.7 Will the data and other research outputs be accessible through a free and standardized access protocol?

- Yes: describe below

RDR provides a free access to the data

2.2.8 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

For sharing restricted information, RDR has the option for asking for specific data. All the data sharing will be done after signing a sharing agreement between both parts.

2.2.9 How will the identity of the person accessing the data be ascertained?

In general, the identity of the persons accessing the data are not necessary. If some sensible data need special attention, the identity of the persons will be controlled by dataset guestbook.

2.2.10 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

- No

No access committe is necessary.

2.2.11 Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why.

- Yes

The data will be done open access under Creative Commons Attribution 4.0 International (CC-BY)

2.2.12 Will metadata contain information to enable the user to access the data?

- Yes

The metadata will provide all the necessary information for finding and accessing the data.

2.2.13 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

The most relevant data will be kept for a minimum of 10 years. In addition, RDR ensures storage for 10 years.
Relevant data includes all protocols required to reproduce the samples for future use, as well as measurements that cannot be repeated due to their cost or the effort involved (i.e. synchrotron measurements).

2.2.14 Will documentation or reference about any software needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

All the characterization techniques will use formats compatible with open source software (.csv).
Some systems (such as fluorescence and FLIM microscopes) use proprietary formats (under license). This software may provide more accurate data analysis, but it can always be converted in .csv. In these cases, the proprietary software will be mentioned, and the .csv file will be uploaded for sharing. The specific format for that software can be share upon request.

2.3 FAIR data: Making data interoperable

2.3.1
What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines?

Will you follow community-endorsed interoperability best practices? Which ones?

All the data and metadata will be transformed into common formats (.csv or tif.). The use of specific software can facilitate the data analysis in certain circumstances. In those cases, the software name will be also provided together a clear explanation of the advantages of that software. The data in proprietary formats can be shared upon request.
Metadata will follow the standard vocabulary used in this research field.

2.3.2 In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies:
Will you provide mappings to more commonly used ontologies?
Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

Standard nomenclature will be used during the project and reports. No project-specific vocabulare will be created.

2.3.3 Will your data and other research outputs include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

- No

2.4 FAIR data: Increase data re-use

2.4.1 How will you provide documentation needed to validate data analysis and facilitate data re-use?

An electronic lab notebook will be created as a txt file. This notebook will contain for each relevant experiment for the research a brief description of the sample used with complete information of synthesis, the parameters used for measuring, and the researcher's observations. The most relevant results will be shared with the corresponding lab notebook to provide all the information.
The experimental data will be recorded with a very descriptive file name (for example: date_material_synthesis conditions) in such a way that all the data can be re-analysed in the future if needed and the samples can be reproduced.
Through this storage protocol, all the data can be recovered, understood, and compared with the researcher's notebooks.

2.4.2
Will your data and other research outputs be made freely available in the public domain to permit the widest re-use possible?
Will your data and other research outputs be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

The outputs of the project will be made open under CC BY licence. Pre-prints will be published immediatly through institutional repository.

2.4.3 Will the data and other research output produced in the project be useable by third parties, in particular after the end of the project?

- Yes

Researchers continuing this line of investigation will be able to use the results of this project as a starting point, as it covers the foundational aspects of the research.
In addition, the proposed approach used in this project can lead to its application in different research areas.

2.4.4 Will the provenance of the data and other research outputs be thoroughly documented using the appropriate standards?

- Yes

2.4.5 Describe all relevant data quality assurance processes.

The quality of measurements will be guaranteed by multiple measurements. Training in different techniques will be provided by experts in each area. The measurement protocol will be optimized to ensure the accuracy and documented as an SOP for future researchers re-using the data.
These data will be validated by experts in the research field, ensuring the soundness of the methodology and interpretation of results.
The data will be published in peer-reviewed journals, ensuring the highest quality.

3. Other research outputs

3.1 Do you have any additional information, that was not addressed in the previous sections, which you wish to provide regarding other research outputs that are generated or re-used throughout the project?

The SOPs generated during the project can be shared on demand taking into account that experimental results in some cases are highly dependent on the lab equipments.

4. Allocation of resources

4.1 What will the costs be for making data and other research outputs FAIR in your project?

Institutional repositories (RDR and LIRIAS) are free of charge. These repositories are included in the IT infrastructure of KU Leuven (host institution of the project).
The costs of journal open access can be estimated round 2000€/publication.

4.2 How will these be covered?

These costs will be covered by the project budget.

4.3 Who will be responsible for data management in your project?

The main responsible of the data management is Rocío Ariza García (postdoctoral researcher of the MSCA project).
In the case some research are done by third parties (collaborators and master/PhD students), she will be the responsible to organize the data with the previously described protocol for the correct use in future.

4.4 How will long term preservation be ensured?

All data published in RDR will be available for at least 10 years. In parallel, the most important results and experimental data related with them will be transferred to a hard disk and digital storage in OneDrive of Roeffaers group. These are managed by the IT department of KU Leuven.
Important results are considered those that have been published in peer-review journals, just like those that the main researcher (Rocío Ariza García) consider of especial relevance for the project continuity.
The responsible person for the long term preservation of the data is Prof. Maarten Roeffaers.

5. Data security

5.1 What provisions are or will be in place for data security?

All the electronics data generated during the project will be stored on local computers (managed by IT department) just as in OneDrive folder.

Regarding the physical data, the samples will be stored in the proper place considering the special conditions required for each type of sample (glovebox, desiccator, regular box). The researcher's notebooks will be accessible in the office of the researchers.

The digital data in local computers must be updated in the OneDrive space weekly. In addition, periodical backups in the KU Leuven network K-drive of the research group will be saved. The TeraCopy Software will be used for theses backups to avoid potential copy fails.

5.2 Will the data be safely stored in trusted repositories for long term preservation and curation?

- Yes
- No

Data shared by RDR is stored for minimum of 10 years. Large volume of data will be preserved in the K-drive of Roeffaers group for the same 10 years.

6. Ethics

6.1 Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?

- No

6.2 Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

- Not applicable

7. Other issues

7.1 Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

- No