DMP title

Project Name DMP - 1S68322N - DMP title

Project Identifier 1S68322N

Principal Investigator / Researcher Nele Merkelbach

Project Data Contact nele.merkelbach@hotmail.com

Description The main objective of this research project is to control particle size, shape and flowability by spherical agglomeration in a pulsating continuous sonsocrystallizer. The objective is seperated in two parts: (1) to investigate the correlation between flowability data and inline microscopy and image analysis data. (2) To visualize the occurring mechanism and investigate the effect of process parameters. A variety of experimental data such as images and videos will be generated and processed to determine particle properties under different experimental conditions.

Institution KU Leuven

1. General Information
Name applicant

Nele Merkelbach

FWO Project Number & Title

1S68322N - Particle size, shape and flowability control by spherical agglomeration in a pulsating continuous sonocrystallizer

Affiliation

- KU Leuven

2. Data description
Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of Data	Format	Volume	How created
1. Design of reactors			
Technical drawings / 3D models of the designed reactors	.ipt / .stp	100 MB	In Inventor Software
2. Chemical experiments (Spherical crystallization experiments) performed in the CIPT lab in Diepenbeek			
Microscopic images	.jpeg	500-800 GB	Optical microscope equipped with a IDS camera
Microscopic videos	.avi	200 GB	Optical microscope equipped with a IDS camera
High speed camera videos	.mraw	1 TB	Photron fastcam mini AX100 equipped with a navitar microlens is used to visualize particles and droplets in flow.
Analyzed images and videos	.txt and converted to Microsoft Excel	16 MB	Analizing the images and videos with ImageJ
Balance measurments	.txt and converted to Microsoft Excel	100 GB	Logging mass changes in function of time during capillary rise tests
Results flowability measurements	Microsoft Excel, pdf, jpeg	500 GB	Granudrum and tapped density measurements
Equation to predict correlation between flowability and particle properties	.txt/ Microsoft Excel	20 MB	Turingbot, correlation between microscopic parameters and flowability
3. Automated routines: created to analyse data (images, videos, correlations,...)			
Image analysis algorithm	.txt	1 MB	Already created in the research group, but changes for this specific application are possible.
Statistical modeling algorithm	.txt	1 MB	Already created in the research group, but changes for this specific application are possible.

3. Legal and ethical issues
Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference:
Short description of the kind of personal data that will be used:

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, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata
What documentation will be provided to enable reuse of the data collected/generated in this project?

1. Microscopic images + videos the following information will be noted: microscope setting, flowcell dimensions, pixel sizes and methodolgy. This info will be summarized in an excel sheet together with the experiment settings (see further).

2. High speed camera videos: the lens, zoom settings, a calibration image, the shutter speed and frame rate will be noted. Additionally, the settings of the light source will also be indicated. The metadata will automatically be saved for each image and this will not be altered.

3. Analyzed images + videos: the analysis gives a txt file which summarises the results and the analysis settings. These files are stored automatically in the folder of the video/image.

4. Algorithm for image analysis: Changes in the algorithm are saved as a new version of the algorithm (txt file) and added to Github. A summarizing txt file is used to document the changes in the different versions of the algorithm.

5. Balance measurements: A summarizing Excel sheet is used to document the settings of the performed cappilary rise tests. This info contains: date and time, experiment number, methodology, type of API, mass of API, sample preparation, type of solvent and volume solvent. The txt-file and excel-sheet of the balance measurement is named after the corresponding experiment number.

6. Crystallization samples: A summarizing Excel sheet is used to document the settings of the performed crystallizatin experiments. This info contains: date and time, experiment number, methodology, description of set-up, set-up settings, type of API, concentration of solution, type, flow rate and volume of anti solvent, type, flow rate and volume of bridging liquid, temperature profile and stirring rate and time.

7. Flowability measurements: A summarizing Excel sheet is used to document the settings of the performed flowability measurements. This info contains: date and time, experiment number, methodology, type, mass and size of API and sample preparation. The excel sheet with results of these measurements are named after the corresponding experiment number.

8. Statistical data analysis algoritm: changes are saved in a new version of the algortihm, back ups are saved on the NAS server of the research group.

A README file will be compiled and updated after every publication. This file will contain a detailed research protocol, a description of the experimental setup, an overview of the investigated conditions, a description of how the data is generated and how it should be

analyzed. This file will allow to re-do experiments and analysis.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

There is no metadata standard available.

5. Data storage and backup during the FWO project
Where will the data be stored?
The research group has a storage capability of around 20TB made by HDD both portable and not. KULeuven has also a One Drive with a storage capability of 2TB per person. The data will be stored
on both during my research and on the HDDs during the 5 years after the end of it. During the research years I will co-manage these data together with prof. Braeken. After the end of my research project, the management responsibility of the data will pass to Prof. Braeken.

- During the research: data will be stored on the personal computer of the researcher, a HDD and on the kuleuven OneDrive account of the researcher. Within 6 months a NAS server will also be available in the research group and used.
- After publication of an article: the data will be stored on the central servers of the university and in the near future also on the NAS.

How is backup of the data provided?

- The kuleuven OneDrive account of the researcher (daily basis)
- An external HDD (monthly basis)
- In the near future data will be stored on the NAS server of CIPT, it will be updated every 3 months.

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 total capicity of the HDD within our research group is around 20 TB and the OneDrive from KU Leuven is 2 TB per person.

What are the expected costs for data storage and back up during the project? How will these costs be covered?
Extra portable HDD can be purchased (150 euros for 5 TB)
The costs can be covered by the general operational budget of the research group CIPT.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?
The computer of the individual researcher is password protected.
The OneDrive account has a two-factor authentication.
The HDD will be with the researcher or in a locked cupboard in the office (office is also locked).
The NAS will only be accesible with unique login credentials.

6. Data preservation after the FWO project
Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained for the expected 5 year period after the end of the project. Redundant microscope and high-speed images and videos will be removed to reduce the size of the stored data. The data will be stored on the university's central servers and the NAS of the research group,
conforming with KU Leuven's RDM policy.

Where will the data be archived (= stored for the longer term)?

In the OneDrive that will be transferred to Prof. Braeken, the NAS or potentially the KU Leuven archive.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?
Extra personal HDD if needed (150 € per 5T)

The KU Leuven Archive costs 169€ per TB per year. It is expected that approximately 5 TB of KU Leuven Archive will be needed (after removal of redundant images). As a result, the total cost over 5 year is estimated at €4300. This will be covered by the general operational budget of Prof. Braeken.

7. Data sharing and reuse
Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?
1. After a publication, all datasets will be made available upon request, to whoever requests the data.
2. Unpublished data will only be accesible to the involved researcher and Prof. Braeken as it may be used in future projects.

Where/how will the data be made available for reuse?

- Upon request by mail

When will the data be made available?

- Upon publication of the research results

After a publication, all datasets will be made available upon request.

Who will be able to access the data and under what conditions?
1. After a publication, all datasets will be made available upon request, to whoever requests the data.
2. Unpublished data will only be accesible to the involved researcher and Prof. Braeken as it may be used in future projects.

What are the expected costs for data sharing? How will the costs be covered?
The costs are those for open access. It will be paid by general working budget of Prof. Braeken.

8. Responsibilities
Who will be responsible for data documentation & metadata?
Nele Merkelbach will be responsible for the implementation of this DMP.

Who will be responsible for data storage & back up during the project?
Nele Merkelbach as well as KU Leuven ICTS managing OneDrive will be responsible for storage and back up of data during the project.

Who will be responsible for ensuring data preservation and reuse ?
Prof. Braeken as well as KU Leuven ICTS managing OneDrive will be responsible for ensuring data preservation and reuse.

Who bears the end responsibility for updating & implementing this DMP?
Prof Braeken bears the end responsibility of updating & implementing this DMP.