FWO DMP Template

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

1. General Information	
Name applicant	Marthe Vanhulst
FWO Project Number & Title	1S47622N - Machine learning based toolpath optimization for multi stage incremental forming.
Affiliation	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other:
2. Data description	
Will you generate/collect new data and/or make use of existing data?	☒ Generate new data ☐ Reuse existing data

Describe the origin, type and format of the data (per dataset) and its (estimated) volume

*If you **reuse** existing data, specify the **source** of these data.*
*Distinguish data **types** (the kind of content) from data **formats** (the technical format).*

The data are collected for the FWO doctoral research project 'Machine learning based toolpath optimization for multi stage incremental forming'.
The project aims to generate an optimal multi-stage toolpath design for Incremental Sheet Forming (ISF) of metal sheets in order to improve accuracy and enhance the process limits. The influence of a multi-stage approach for enhancing the process limits and broadening the application range of Single Point Incremental Forming (SPIF) is investigated. A Machine Learning algorithm will be implemented and used for predicting a multi-stage process outcome. This will enable finding a multi-stage approach that optimizes the final shape for thickness distributions and accuracy.

The following table gives an overview of the collected data, its type, format and estimated volume:

Type of data	**Format**	**Volume**	**How created**
Data of parts to be manufactured (CAD models, toolpaths)	.stl, .par, .mxp, .cls, .ptp	5 GB	Siemens NX, Siemens Solid Edge, Materialise 3-Matic, JSPIF (in-house processing software)
Manufactured parts	Hardware: sheet metal (Aluminium alloys, Sink)	~ 50 samples of 225x225 mm	Manufactured with Single Point Incremental Forming using a KUKA robotic arm
Raw measuring data (geometries, thickness, DIC images, …)	GOM data: .gscan, .ginspect, MatchID data: .tiff, .cal, .calfat .3dat, .mti3d	1 TB	Measured data during and after production with fringe projection (GOM Athos Compact Scan) and Digital image correlation with MatchID Stereo DIC.
Processed measuring data (geometries, numerical data, …)	.stl, .csv, .png	50 GB	Post-processed in GOM inspect, MatchID Stereo and Python code.

	Source files for data processing	.py, .cfg	2 GB	Developed Python code (using Pycharm) for processing the data and applying Machine Learning. Stored in GitLab.
	Datasets with relevant Machine Learning parameters	.csv, .txt	1 GB	Developed with data post-processing code to extract the important parameters for ML.
	Metadata describing strategies and setup for collecting and processing data + documentation	.docx, .txt, .xlsx	< 1 GB	Microsoft Office Word and Excel, Notepad.
	Scientific output data: written papers, presentations, figures, etc.	.docx, .ptp, .png, .jpeg	< 1 GB	Microsoft Office Word and Powerpoint, results from measurements and from post-processing scripts.

3. Ethical and legal issues

Will you use personal data? If so, shortly describe the kind of personal data you will use AND add the reference to your file in your host institution's privacy register. *In case your host institution does not (yet) have a privacy register, a reference is not yet required of course; please add the reference once the privacy register is in place in your host institution.*	☐ Yes ☒ No If yes: - 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).	☐ Yes ☒ No If yes: - Reference to ethical committee approval:
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?	☒ Yes ☐ No If yes, please comment: Experiments used for publications will be made available on a database that is under development within the ISF research group at RWTH Aachen, in collaboration with our research group (Flexible Sheet Metal Working group at KU Leuven). This enables tech transfer to other interested research groups, as well as a transparent availability of reported results.
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?	☐ Yes ☒ No If yes, please comment:

4. Documentation and metadata

What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?	**Source files for data processing** The Python project containing all source files for the data processing will be accompanied by a ReadMe file and all necessary documentation in function descriptions and as in-line comments. Typing will be used to clarify and check input arguments. The code will be made easily transferable by making packages that can be installed by other users. Version control will be handled with GitLab. **Experimental data** The data for each experiment will be stored in a different folder with a fixed hierarchical structure. Naming conventions will be used to ensure easy searching. Versions will be tracked using version control of OneDrive and autobackup. The structure of an experimental data folder looks as follows: • CAD models: for the construction of the parts (mxp, prt, stl, …) • Toolpaths: the toolpaths used for each part (cls, ptp, …) • DIC measurements: - calibration (.tiff, .cal, .caldat) - images taken during the experiment (.tiff) - correlation files (mti3d) and corresponding correlation results (.3dat) • GOM measurements: - Measurements of the produced parts (.stl, .gscan, .ginspect) • Results: - Results processed with python: figures (.png), dataframes (.csv) - Results summary and figures (.svg, .ppt, .docx, …) **Descriptions of setup and strategies** An experiment might be accompanied by some extra ReadMe files and documents explaining the strategy, experimental setup and variables used for processing the data. This will be described in separate Microsoft Word documents.

Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.	☐ Yes ☒ No If yes, please specify: No metadata standard has been in use for the Incremental Sheet Forming process. The collaboration with the ISF group of RWTH Aachen might include the development of a new metadata standard for the ISF process, but enrolling this feature will take up time. Meanwhile, during the research, the data will be structured the same way for all experiments with naming conventions that are clear and include the same pre- and/or suffix to avoid confusion. Additional documents that contain the parameters and workflow for each experiment will be stored in the accompanying folder.

5. Data storage & backup during the FWO project	
Where will the data be stored?	All data described above will be stored on the personal KU Leuven OneDrive account.
How will the data be backed up?	The data on the personal KU Leuven OneDrive account is automatically backed up. On top of that, the data will also be backed up on an external hard drive on a weekly basis.
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 ☐ No If no, please specify: The total volume of data generated and stored during the research is currently at 500 GB and is estimated to be at 2-3 TB by the end of the research. The storage of the raw measured data takes up the most storage space, since DIC measurements contain hundreds of images and CSV-files per experiment, as well as 10-20 STL-files representing the CAD models and measured shapes after the production of each step. This is a high amount of space needed and will probably exceed the available capacity of the OneDrive account of 2 TB provided by KU Leuven by the end of the research. During the research, data from previous work packages or finished papers will be transferred to a hard drive of 2 TB to ensure sufficient capacity at all times. The cloud storage will be increased to ensure a back-up of all experimental data.

What are the expected costs for data storage and backup during the project? How will these costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of* ***the allocated project budget*** *to be used to cover the cost incurred.*	The prices of an external hard drive of 3TB are currently around 90-110 euros, which will be enough storage capacity for the entire research project. This amount may still decrease in the upcoming years due to More's law resulting in more efficient and cheaper storage. The cloud capacity might need to be increased in the final years of the research, when the current limit of 2 TB might be exceeded. Storage of 1 TB currently costs between 7 and 9 euros per month, resulting in an extra cost of around 100 euros per year. This extra storage will probably only be needed at the end of the research, and is estimated to be 100-300 euros in total. The total costs for cloud and hard drive storage are estimated at 240-440 euros and will be covered with an internal operational budget.
Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?	The cloud storage is only accessible by authorized persons. Hard drives will be stored in a locked room.

6. Data preservation after the end of the FWO project FWO expects that data generated during the project are retained for a period of minimally 5 years after the end of the project, in as far as legal and contractual agreements allow.	
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, ...).	After the end of the employment, the personal KU Leuven OneDrive storage space will no longer be accessible. All data will then be accessible on the purchased hard drives to allow availability of the data for the expected time period of 5 years.
Where will these data be archived (= stored for the long term)?	On a hard drive.

What are the expected costs for data preservation during these 5 years? How will the costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of* ***the allocated project budget*** *to be used to cover the cost incurred.*	The prices of an external hard drive of 3TB are currently around 90-110 euros, but this price can significantly reduce in the upcoming three years. Depending on the storage needed, one or two hard drives can be purchased. The costs will be covered with internal budget.

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)?	☐ Yes ☒ No If yes, please specify:
Which data will be made available after the end of the project?	Datasets including the processed measured data, as well as the pre-processed data needed to manufacture the parts. These data have large added value, since it enables reproduction and shares knowledge on the process behaviour and outcomes.
Where/how will the data be made available for reuse?	☒ In an Open Access repository ☐ In a restricted access repository ☐ Upon request by mail ☐ Other (specify): Datasets used in publications may be made available on an Open Access database which is currently under construction in cooperation with RTWH Aachen as a platform to share data on Incremental Sheet Forming.
When will the data be made available?	The data might be shared after publication of the corresponding article or conference proceedings.

Who will be able to access the data and under what conditions?	The data will be accessible to all people interested in ISF that have an account on the shared platform. This database platform is still under development, so no large community has been reached yet. After finalising the database and a trial run, it will become a broadly available community for all groups interested in ISF.
What are the expected costs for data sharing? How will these costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of **the allocated project budget** to be used to cover the cost incurred.*	The database will be funded by internal budget of the ISF research group at RTWH Aachen. A funding application will be submitted for the ESAFORM benchmark grant in September 2022.

8. Responsibilities	
Who will be responsible for the data documentation & metadata?	Marthe Vanhulst
Who will be responsible for data storage & back up during the project?	Marthe Vanhulst
Who will be responsible for ensuring data preservation and sharing?	Marthe Vanhulst
Who bears the end responsibility for updating & implementing this DMP? *Default response: The PI bears the overall responsibility for updating & implementing this DMP*	Marthe Vanhulst