InSituPro DMP

InSituPro Data management Plan

ADMIN DETAILS

Project Name: InSituPro DMP – InSituPro Data Management Plan
Principal Investigator / Researcher: Yentl Swolfs
Project Data Contact: +32 16373616, yentl.swolfs@kuleuven.be
Description: The key objective of InSituPro is to unravel the constituent properties of fibre-reinforced polymer composites (FRPCs). To achieve this goal, we set up a modelling-supported experiments approach. Our methodology will integrate digital volume correlation (DVC) into synchrotron radiation computed tomography (SRCT) tomographs, to obtain 3-dimensional strain maps that indicate/monitor the *in-situ* damage behaviour of FRPCs. In this methodology, each unit operation requires considerable storage capacity. More specifically, raw image acquisition, image processing and analysis, digital volume correlation, finite element (FE) verification models and *in-situ* mechanical testing data could be referred to as the building blocks for the set objective.
Institution: KU Leuven

1. GENERAL INFORMATION

Name of the project lead (PI)
Yentl Swolfs

C1-C2 Project number & title
C14/21/076 Unravelling *in-situ* constituent properties in fibre-reinforced composites: digital volume correlation and machine learning applied to synchrotron computed tomography images

2. DATA DESCRIPTION

2.1. Will you generate/collect new data and/or make use of existing data?
- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

	Format	Volume	How created

Type of Data			
X-ray computed tomography volumes	.tif	2 TB	XCT core facility – KU Leuven
Synchrotron computed tomography images	.tif	20 TB	Synchrotron radiation facilities (European Synchrotron Radiaton Facility, Swiss Light Source)
DVC fields	.am, .csv	2 TB	Avizo Software (Thermo Fisher Scientific)
Optical Images	.tif	100 GB	*In-situ* using optical cameras for mesoscale DIC

3. ETHICAL AND LEGAL ISSUES

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.
No

3.2. 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

3.3. Does your research 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

3.4. 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 regarding reuse and sharing are in place?
No

4. DOCUMENTATION AND METADATA

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

1. <u>For the X-ray computed tomography volumes</u>: Sample type, layup and dimensions, source and detector specifications, load and displacement for each *in-situ* scan, scan settings (voltage, current, exposure time, number of projections, number of averaging procedures, voxel size, source to object distance, source to detector distance).

2. <u>For the synchrotron computed tomography volumes</u>: Sample type, layup and dimensions, source and detector specifications, load and displacement for each *in-situ* scan, scan settings (voltage, projections, exposure), propagation distance, beam type, camera specifications.

3. <u>For the DVC fields</u>: Sample type and dimensions, DVC approach, sub-volume size, load/displacement step.

4. <u>For the optical images</u>: Sample type, layup and dimensions, camera specifications, image acquisition rate, region of interest, subset size, step size.

4.2. 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.

Metadata generation is performed automatically for the computed tomography operations (1) and (2), e.g. generation of the scan settings.txt file through the CT acquisition software.

For the DVC fields and the optical images, manual metadata will be created by the researchers, incorporating the data stated in section 4.1.

For CSV and TIFF file extensions, we follow the ISA-Tab standard.

For the CT data, we follow the NeXus standard.

5. DATA STORAGE AND BACKUP DURING THE C1-C2 PROJECT

5.1. Where will the data be stored?

For all data, we will use external hard disk drives (HDDs) while creating a data backup on our research group NAS system.

5.2. How will the data be backed up?

Since the start of the project (01/10/21), we have purchased two HDDs for the exclusive purpose of backing up the generated data, in addition to a full backup on our NAS system. Furthermore, selected data are stored in our research group (Composite Materials Group) shared folder and in the PI's personal OneDrive folder.

5.3. 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, there is enough space on our HDDs, NAS, shared folder and OneDrive personal folder. The available storage is expected to be sufficient to include the data generated within the four years of this doctoral study. The purchase of an additional HDD could be an option in the future but is judged as non-essential at this moment.

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

We primarily expect a moderate cost of purchasing HDDs.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The PI has the HDDs under his possession. He could allow access to selected individuals if desired.
Only the PI has access to his personal OneDrive folder.
Only authorised personnel (within our group) have access to our local NAS system and shared drive. This system is password protected.

6. DATA PRESERVATION AFTER THE END OF THE C1-C2 PROJECT

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained for the expected 10 year period after the end of the project.

6.2. Where will these data be archived (= stored for the long term)?

External HDDs and NAS.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

No costs are anticipated (excluding the moderate cost of HDD replacements).

7. DATA SHARING AND RE-USE

7.1. 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 or because of IP potential)?

No

7.2. Which data will be made available after the end of the project?

While the complete dataset is anticipated to be too large for open access availability, exemplary sub-datasets are expected to be made available after the end of the project or/and upon publication of the research results.

7.3. Where/how will the data be made available for reuse?
- In an Open Access repository. We anticipate storing datasets on Mendeley Data and Data in Brief, both of which have been extensively used from our research group in the past.

7.4. When will the data be made available?
- Upon publication of the research results. Furthermore, currently synchrotron facilities make all datasets from beamtimes publicly available.

7.5. Who will be able to access the data and under what conditions?
Upon publication of the research results, these will become available to the public (e.g. Mendeley Data).

7.6. What are the expected costs for data sharing? How will these costs be covered?
Mendeley Data provides 10 GB of free space which should be sufficient for our purposes.
Any anticipated costs related to data sharing and publications (e.g. Data in Brief article publishing charge amounts to 500 USD, excluding taxes) will be paid from the project working costs.

8. RESPONSIBILITIES

8.1. Who will be responsible for the data documentation & metadata?
Thanasis Chatziathanasiou

8.2. Who will be responsible for data storage & back up during the project?
Thanasis Chatziathanasiou

8.3. Who will be responsible for ensuring data preservation and sharing?
Thanasis Chatziathanasiou together with Prof. Yentl Swolfs

8.4. Who bears the end responsibility for updating & implementing this DMP?
Prof. Yentl Swolfs