Call: HORIZON-WIDERA-2021-ACCESS-03/Twinning

Project SustDesignTex (GA No. 101079009), title: „Sustainable Industrial Design of Textile Structures for Composites" funded by the European Union

The first version of the Data Management Plan

SUSTainable industrial DESIGN of TEXtile structures for composites (SustDesignTex)

Grant Agreement number: 101079009

Project acronym: SustDesignTex

Project title: Sustainable Industrial Design of Textile Structures for Composites

Funding Scheme: HORIZON-WIDERA-2021-ACCESS-03

Start date of the project: 01.10.2022

Duration: 36 months

Project coordinator name, title, and organization: Marcin Barburski, DSc, prof. TUL, Lodz University of Technology

Project coordinator e-mail: marcin.barburski@p.lodz.pl

Project e-mail: sustdesigntex@info.p.lodz.pl

Document History (Revisions – Amendments)		
Version	**Date**	**Changes**
V 1.0	17.03.2023	First version
V 1.1	20.03.2023	Tsegaye Lemmi's changes

Main Author of the Deliverable, Consortium Beneficiary	Other Contributors, Consortium Beneficiary
TUL	-

Call: HORIZON-WIDERA-2021-ACCESS-03/Twinning

Project SustDesignTex (GA No. 101079009), title: „Sustainable Industrial Design of Textile Structures for Composites" funded by the European Union

Deliverable Title	**The first version of the Data Management Plan (DMP)**
Milestone Lead:	TUL
Related Work Package:	Work package WP8 – Project management and coordination
Author(s):	Agata Poniecka
Due Submission Date:	31/03/2023
Actual Submission:	31/03/2023

Table of content

1 DESCRIPTION OF DATA – RE-USE OF EXISTING DATA AND/OR PRODUCTION OF NEW DATA ..3

2 DOCUMENTATION AND DATA QUALITY ..3

3 STORAGE AND BACKUP ..4

4 LEGAL AND ETHICAL ASPECTS ..4

5 ACCESSIBILITY AND LONG-TERM STORAGE ..5

6 RESPONSIBILITY AND RESOURCES ..6

1 Description of data – re-use of existing data and/or production of new data

• What types of data will be created and/or collected, in terms of data format and amount/volume of data?

The project will produce the following type of data:

(1) Stakeholder contacts collection: publicly available data;

(2) Textile Reinforcement Composites collection: publicly available data;

(3) Textile industrial design of innovative textile structures for composites solutions collection: primary data;

(4) Quantitative experimental data: primary data; the data contains quantitative information on the innovative textile structures for composites at different stages of their lifecycle;

(5) Workshops, trainings, the Joint Summer School data: primary data; the data contain protocols, written notes, and summaries that were done at the workshops, trainings, and the Joint Summer School, which will be organized according to the project Work Programme;

(6) Validation cycles data: primary data.

There will be used Microsoft Office for text-based documents (or any other compatible version) .doc, .docx, .xls, .xlsx, .ppt, .pptx. Also, especially where larger datasets need to be dealt with, .csv and .txt file formats will be used. All finished and approved documents will also be made available as .pdf documents. Illustrations and graphic design will make use of Microsoft Visio (Format: .vsd), Photoshop (Format: different types possible, mostly .png), and will be made available as .jpg, .psd, .tiff, and .ai files. PFDs, PIDs and layouts will preferentially use inkscape.org, an open-source vector graphics software (Format: .svg), and will be made available as .png, .jpg, and .pdf files. MP3 or WAV will be used for audio files and Quicktime Movie or Windows Media Video – for video files. These file formats have been chosen because they are accepted standards and widely used. Files will be converted to open file formats where possible for long-term storage.

• How will data be collected, created, or re-used?

The data collected will mainly come from the tests carried out. The data collection format will depend on the testing software. The data will be collected and stored on the pendrives of the employees responsible for them, specially created for the cloud drive project, and on a backup copy on a physical drive.

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

There is an expected amount of a few gigabytes of generated data.

2 Documentation and data quality

• How will the material be documented and described, with associated metadata relating to structure, standards, and format for descriptions of the content, collection method, etc.?

In order to be able to distinguish and easily identify data sets, each data set will be assigned with a unique name. This name can also be used as the identifier of the data sets. All data files produced, including emails, include the term "SustDesignTex", followed by a file name that briefly describes its content, followed by a version number (or the term ''draft" or "FINAL"), followed by the short name of the organization which

prepared the document (if relevant). Each data set that will be collected, processed, or generated within the project will be accompanied by a brief description.

• How will data quality be safeguarded and documented (for example, repeated measurements, validation of data input, etc.)?

Data quality can be safeguarded and documented through various methods, including:

- Repeated measurements: One of the best ways to ensure data quality is to take multiple measurements of the same parameter or variable. This helps to identify any errors or anomalies and allows for verification of the data.
- Validation of data input: Before data is entered into a system or database, it should be validated to ensure that it meets certain criteria. For example, data can be checked for completeness, consistency, accuracy, and validity.
- Standardized data collection protocols: Data collection protocols should be standardized to ensure that data is collected consistently and accurately. This can include using standardized forms, procedures, and techniques.
- Quality control checks: Quality control checks can be used to ensure that data is accurate and complete. These checks can include things like data entry verification, data cleaning, and data validation.
- Documentation: All data should be documented thoroughly, including the date and time of collection, who collected the data, any relevant environmental conditions, and any other pertinent information. This documentation can help to identify any issues with the data and can be used to support data analysis and interpretation.

3 Storage and backup

• How is storage and backup of data and metadata safeguarded during the research process?

The data collected during the project is stored on a cloud drive to which representatives of individual partners have access. Employees can only access the drive from their work accounts. Every two weeks, a copy of this disk is made, and the data is ripped to the physical disk.

• How data security and controlled access to data safeguarded, in relation to the handling of sensitive data and personal data, for example?

Access to sensitive data is available only to the necessary persons - the project manager and persons dealing with the financial and administrative service of the project. At TUL, these data are stored and processed in accordance with the procedures in force at the university and dedicated tools are used for this purpose.

4 Legal and ethical aspects

• How is data handling according to legal requirements safeguarded, e.g. in terms of handling of personal data, confidentiality and intellectual property rights?

Data handling according to legal requirements is safeguarded by adhering to applicable laws and regulations regarding data protection, confidentiality, and intellectual property rights. This includes:

1. Handling of personal data: If the research involves handling of personal data, researchers must comply with relevant data protection laws and regulations, such as the General Data Protection Regulation (GDPR) in the European Union.This includes obtaining informed consent from participants, ensuring that data are kept confidential, and providing participants with the right to access and control their data.
2. Confidentiality: Researchers must take appropriate measures to ensure the confidentiality of research data, particularly if the data contain sensitive or identifiable information. This may involve using secure storage and transmission methods, implementing access controls, and de-identifying data before sharing them. At the moment, the project does not produce confidential research results. Confidential data includes only personal data related to employment. This data is stored and processed in accordance with the internal regulations of each partner.
3. Intellectual property rights: Researchers must respect intellectual property rights, such as patents, copyrights, and trademarks, when handling research data. This includes obtaining necessary permissions and licenses for using or sharing data, and properly attributing sources of data. According to the Consortium Agreement point 8.1: *Results are owned by the Party that generates them.*

To safeguard data handling according to legal requirements, researchers may need to consult with legal experts and data protection officers to ensure compliance with applicable laws and regulations.

• How is correct data handling according to ethical aspects safeguarded?

During the project realization, no research involving humans or animals will be conducted. Therefore, the project is not concerned with the aspect of data ethics.

5 Accessibility and long-term storage

• How, when, and where will research data or information about data (metadata) be made accessible? Are there any conditions, embargoes, and limitations on the access to and re-use of data to be considered?

Researchers may choose to deposit their data in open-access repositories or data archives, such as Dryad or Zenodo, or in institutional repositories, such as those maintained by universities or research organizations. The data will be made available, first and foremost, via the TUL Digital Repository of Science CYRENA (CYfrowe REpozytorium NAuki: http://repozytorium.p.lodz.pl/?locale-attribute=en), which collects, archives and makes available the academic achievements of TUL staff in the full-text form. The repository contains primarily scientific articles, books, monographs, conference materials, presentations, and data storage. In some cases, data may be deposited in subject-specific repositories that cater to a particular community of researchers.

The open research data will also be archived on the Zenodo platform (http://zenodo.org), an EU-backed portal based on the well-established GIT version control system (https://git-scm.com) and the Digital Object Identifier (DOI) system (http://www.doi.org).

Metadata should also be made accessible, ideally through standardized metadata schemas that are widely accepted within the research community. This metadata should provide sufficient information to enable the data to be discovered, understood, and re-used by other researchers.

The accessibility of research data and metadata will be carefully planned and communicated, taking into account any conditions, embargoes, or limitations on access and re-use. Researchers should aim to make their data and metadata as accessible as possible while also ensuring that any sensitive or confidential information is properly protected.

6 Responsibility and resources

• Who is responsible for data management and (possibly) supports the work with this while the research project is in progress? Who is responsible for data management, ongoing management and long-term storage after the research project has ended?

Each SustDesignTex project partner is responsible for storing data they generate during the realization of the project and providing it to the Lodz University of Technology (TUL) at the time of reporting and whenever requested by the project coordinator institution(TUL). The project coordinator institution (TUL) is responsible for storing the data during and beyond the project's lifespan. According to the Consortium Agreement point 8.1: *Results are owned by the Party that generates them.* It, therefore, requires data collection by the respective partner.

• What resources (costs, labour input or other) will be required for data management (including storage, back-up, provision of access, and processing for long-term storage)? What resources will be needed to ensure that data fulfills the FAIR principles?

The resources required for data management will depend on the scope and scale of the data being managed, as well as the specific requirements for storage, backup, access provision, and processing. At the moment, we only anticipate costs related to the purchase of hard drives for backing up data. During the project implementation, however, costs may arise associated with hardware (e.g., servers, storage devices), software (e.g., data management systems, backup and recovery software), and labor (e.g., IT staff, data curators).

To ensure that data fulfill the FAIR principles, additional resources may be necessary, including:

1. Metadata creation: FAIR data requires rich, accurate, and standardized metadata to enable data discovery and re-use. This may require additional labor and expertise to create and maintain metadata.
2. Data formatting and standardization: Data may need to be formatted and standardized according to community-accepted standards to ensure interoperability and re-use.
3. Data quality control: FAIR data should be of sufficient quality for re-use, which may require additional resources for data cleaning and quality control.
4. Data access and sharing: FAIR data must be accessible and shared in a way that maximizes its re-use potential. This may require additional resources for developing data sharing agreements, providing access to data, and ensuring proper attribution and citation.
5. Long-term storage and preservation: FAIR data must be preserved and accessible over the long term, which may require additional resources for archiving and preservation.

Overall, ensuring that data fulfill the FAIR principles may require a significant investment of resources, including hardware, software, and labor, as well as expertise in data management, metadata creation, data standardization, data sharing, and preservation.