

FWO DMP

FWO DMP Jan De Troij

ADMIN DETAILS

Project Name: *FWO SB fellowship - FWO DMP Jan De Troij*

Project Identifier: *1S32622N*

Grant Title: *1S32622N - Development and deployment of physics-based digital twins for the analysis of the vibro-acoustic behavior of systems of industrial complexity*

Principal Investigator / Researcher: *Wim Desmet*

Institution: *KU Leuven*

1. GENERAL INFORMATION

Name applicant

Jan De Troij

FWO Project Number & Title

1S32622N - Development and deployment of physics-based digital twins for the analysis of the vibro-acoustic behavior of systems of industrial complexity

Affiliation

KU Leuven

2. DATA DESCRIPTION

Will you generate/collect new data and/or make use of existing data?

Research data are any information collected or generated for the purpose of analysis, in order to generate or validate scientific claims. It includes digital and physical data.

Research data encompass the whole spectrum ranging from raw data to the processed and analysed data.

Examples include survey results, statistics, measurements, notebooks, images, texts, computer generated data, simulations, software developed for research purposes, computational metadata, prints, video- and audiotapes, coding of textual information, organisms, gene sequences, synthetic compounds, samples, patients data, etc.

Existing data are all data which you did not generate yourself in the context of this research project but plan to use for the research. This can include online datasets, datasets received upon request or under a data transfer agreement, or existing datasets available within your research unit.

✓ *Generate new data*

~~*Reuse existing data*~~

Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) and per WP or objective of the project.

The data description forms the basis of your entire DMP, so make sure it is detailed and complete. Include:

- all existing data you will use
- all new data your research will generate
- physical samples or analogue data
- input and output data
- analysis scripts and code

This list can be updated as your research progresses.

If you **reuse** existing data, specify the **source** of these data, with citation or weblink and copyright / licence information.

Generated experiment-related data

Type of data	Format	Volume	How created
Raw and processed acoustic and vibration measurements (numerical)	Databases (.lms, .mat, .unv), graphs (.fig)	10-100 GB	In-lab vibration & acoustic measurements using Siemens Test.Lab and Polytec software
Images and videos (multimedia)	Image (.png), video (.mp4)	10-100 GB	Camera images/videos of samples and test setups
Metadata describing measurements setup and procedures (textual)	Text (.txt, .docx, .pptx, .pdf)	< 1 GB	Notepad, Microsoft Office Word, Microsoft Office PowerPoint

Generated simulation-related data

Type of data	Format	Volume	How created
Vibro-acoustic finite element simulation models (model)	Commercial finite element software specific formats (.sim, .fem, .prt, .mph), model matrices (.mat), result files (.pch, .op4, .mat, .unv)	100-300 GB	Siemens Simcenter 3D, Siemens NX, COMSOL Multiphysics, Matlab
Raw and processed simulation data (numerical)	Databases(.mat), graphs (.fig)	< 100 GB	Matlab
Images and animations (multimedia)	Image (.png), video (.gif)	10 – 100 GB	Matlab
Metadata describing models and simulations setup and procedures (textual)	Text (.txt, .docx, .pptx, .pdf)	< 1 GB	Notepad, Microsoft Office Word, Microsoft Office PowerPoint

Reused experiment-related data

/

Reused simulation-related data

/

Dissemination-related data

Type of data	Format	Volume	How created
<i>Paper manuscripts and abstracts (textual)</i>	<i>Text (.pdf, .tex, .docx)</i>	<i>< 100 MB</i>	<i>E.g. Microsoft Office Word, TextStudio (LaTeX)</i>
<i>Presentations and slideshows (textual)</i>	<i>Text (.pptx, .pdf)</i>	<i>< 1 GB</i>	<i>E.g. Microsoft Office PowerPoint</i>
<i>Demonstrator movies (multimedia)</i>	<i>Multimedia (.gif, .mp4)</i>	<i>< 1 GB</i>	<i>E.g. Hitfilm Express, Microsoft Office PowerPoint, Windows Video Editor</i>

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 Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

✓ No

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?

Do your data contain trade secrets, is it part of a patentable invention or viable for commercial exploitation?

Describe which restrictions there may be on future use and sharing of the data. Consider which licence would be best suited for future reuse of the data. This [Licence Selector Tool](#) can help. Note that you can only publish data that you are the owner of.

Consult the KU Leuven [Intellectual Property](#) guidance.

Useful guidance on copyright and IP is also available from [CLARIN](#).

✓ Yes

The work in this project can result in research data which has potential for tech transfer and valorisation:

- *Novel models and methodologies will be developed which interact with and/or are based upon existing models and methodologies of the hosting KU Leuven Division LMSD. The IP of the latter lies with KU Leuven and the Division LMSD and will also hold for the further developments made in this project.*

The KU Leuven Division LMSD has expertise in and an excellent track record regarding tech transfer and valorisation which will be leveraged upon in this project. The research manager of the Division LMSD Bert Pluymers will be consulted regarding these IP aspects.

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?

If you indicated in section 2 that you reuse existing data, check under which conditions or licence reuse is allowed or which restrictions on wider data sharing might be in place. Can you publish the data? Can you share them with other researchers?

If in your research you use data or materials obtained via a DTA (data transfer agreement) or MTA (material transfer agreement), then check which restrictions may apply.

✓ No

It is noted that novel models and methodologies will be developed, which interact with and/or are based upon existing models and methodologies of the hosting KU Leuven Division LMSD research group. In addition, the work in this project fits in the research of the hosting KU Leuven Division LMSD, where related activities within other projects are developed and can leverage upon the developments in this project to enhance interaction. The IP of the models and methodologies of the hosting research group lies with KU Leuven and the Division LMSD and will also hold for the further developments made in this project and for future developments based on these. Reference to existing models and methodologies of the hosting KU Leuven Division LMSD will be made by citing the associated journal paper publications.

4. DOCUMENTATION AND METADATA

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

Good documentation of your datasets improves the quality of your research data, is essential for future use and reuse of the data, and also prevents misinterpretation of the data. This requires clear data description, annotation, contextual information and documentation that explains how data were created, what the data mean, what their content and structure is, and any processing that may have taken place.

Documentation is important at the level of a dataset (e.g. describing how the data were created), but also at the level of individual data elements (e.g. explaining what each variable means or the parameters for generation of datafiles such as images).

Consider how and where data documentation is easiest captured, for example within the data file or database, in a ReadMe text file, in file headers, in a codebook or lab notebook.

Tools such as [GitLab KU Leuven](#) and the [Open Science Framework](#) can be used to document data processing and have the benefit of also providing good version control of data files.

State-of-the-art data documentation would be to publish a data descriptor describing and promoting your dataset in a [peer-reviewed data journal](#).

Creating comprehensive data documentation is easiest when begun at the onset of a project and continued throughout the research.

Reuse of the data not only refers to reuse by other researchers, but also by yourself in the future, or for replications.

Metadata on experiment-related data

- *Raw vibration and acoustic measurements: a measurement report per measurement dataset will detail the software parameters & instruments settings, dimensions, measurement methodology and procedural information on how the data was collected, required sensor & exciter labels and positions as well as units of measurements and calibration settings.*
- *Processed vibration and acoustic measurements: a readme file will accompany the dataset, describing the labels and definitions of variables, the units of measurements and how the raw data have been processed.*
- *Photos of the measurement setup with clarifying file names, date and timestamps will be added as additional clarification.*

Metadata on simulation-related data

- *Vibro-acoustic finite element simulation models and raw & processed simulation data: a readme file will describe the parameters, definitions, units and software settings used to construct the models/ data. Modelling assumptions and underlying equations will be clearly reported such that each simulation result is accompanied with a description of the validity range of the model/ data. This information is kept either in the CAE file format, or in an accompanying text file.*

In addition to the metadata per datatype as outlined above, a readme file will be created which describes the structure of the dataset.

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

Although these are not formalized, state-of-practice standards will be used regarding metadata. On the one hand, standardized CAE file format information and structure are used. On the other hand, when

working with transferable neutral text formats for experimental and simulation data, common engineering practice is used, deploying tabulated structures with clear column and row headers.

A clear folder structure will be adopted for the data storage, in accordance with the different tasks carried out during the project. In every (sub)directory, a readme file will list all the present subdirectories and files as well as where the data is used and stored. Whenever publication or sharing of a dataset would be considered, it will be re-evaluated if an applicable metadata standard is available and can be applied to enhance sharing.

5. DATA STORAGE AND BACKUP DURING THE FWO PROJECT

Where will the data be stored?

Physical samples (hardware) will be stored in dedicated lab storage of the KU Leuven Division LMSD.

Data will be stored using cloud- and ICTS-based solutions:

- *Local desktop file storage, with regular backups on the researcher's personal KU Leuven network drive.*
- *Regular snapshots on personal external hard drives.*
- *Cloud-based storage, synchronized with local desktop storage: KU Leuven OneDrive, which provides version-control.*
- *GitLab repository of the KU Leuven Division LMSD.*

For specific storage solutions, there is support from ICTS as well as local IT from the KU Leuven science, engineering and technology group.

How is backup of the data provided?

The cloud-based and ICTS storage solutions are backed-up as part of the offered services. Back-ups are provided on different levels:

- *For the data stored on the KU Leuven central servers, automatic daily back-up procedures apply.*
- *KU Leuven OneDrive and GitLab provide automated backups.*
- *Backups from local desktop file storage to personal KU Leuven network drive will be regularly performed (e.g. using SyncBackFree backup software).*

In addition, back-up copies on personal external hard drives will allow to recover data files.

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

Currently sufficient storage & backup capacity is available during the project for the anticipated data volumes. The available storage space and file size limits exceed the currently estimated required storage space. Furthermore, network drive and cloud storage space can be readily expanded upon request to KU Leuven IT services.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

In case expenses are needed, part of the allocated FWO project budget can be used. However, this is currently not expected.

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

On the one hand, in this project no sensitive personal data will be used. On the other hand, both cloud and ICTS based storage solutions are only accessible via proper credentials which are centrally managed. Hence, unauthorized access via legal means is not possible.

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, ...).

The physical samples and demonstrators will be stored in the KU Leuven Division LMMSD's lab storage. In case space limitations would require keeping only a selection of the manufactured samples, a selection will be based on which samples are more easily & low cost reproducible versus which samples are not as well as the required storage space, while prioritizing on the preservation of the demonstrators which have high valorisation and outreach potential.

Particular focus for non-hardware data preservation will be on:

- *Data at the basis of publications such as journal papers, conference papers and presentations or posters. All data related to Open Access publications will be retained on Public Repositories (like Zenodo or Lirias).*
- *Developed models and measurement datasets which are likely to be reused in the research unit for future research and/or valorisation activities and for future research of the researcher. Potentially large (intermediate and non-postprocessed raw) result and measurement files will be discarded to reduce required storage space if the simulation models and experiments allow to recalculate/remeasure the results easily and at low cost and time.*

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

Hardware storage will be done in the foreseen storage space in the KU Leuven Division LMMSD's labs.

Data will be archived on internal KU Leuven data storage facilities. The data will be stored on the university's central servers on a data archive drive (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

In addition, measurement data can be published in data papers describing and promoting the dataset. Published manuscript preprints will be stored on the KU Leuven Open Access repository Lirias. All other long term stored data will be on KU Leuven ICTS servers.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

Expected costs for storage beyond project duration will be limited and covered by the research group.

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)?

If you indicated in section 2 that you reuse existing data, you should check under which conditions reuse is allowed and whether there are restrictions regarding data sharing. Sometimes data are licensed to be reused for research, but not to be distributed further or not to be shared with other people.

If you generate data that have potential for tech transfer and valorisation, be aware that the university holds the IPR.

Consider strategies to limit restrictions on sharing and reuse, such as anonymising or pseudonymising, gaining participant consent for data sharing, gaining copyright permissions or agreeing a limited embargo period.

No.

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

For the datasets listed in section 6, specify which will be available for future use. If no data can be made available to others, say so explicitly and explain why this is the case.

If some data cannot be shared, or only under restrictions, provide a rationale for doing so, clearly separating legal and contractual reasons from voluntary restrictions.

Physical samples and demonstrators will be made available after the end of the project.

Already throughout the course of the project, all data (measurements, simulations, models) related to the Open Access publications will be made publicly available. After the end of the project, consolidated datasets linked to validation cases will be made available as well.

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):~~

All relevant datasets will be made available via Zenodo or similar.

When will the data be made available?

✓ *Immediately after the end of the project*

✓ *Upon publication of the research results*

~~— *After an embargo period. Specify the length of the embargo and why this is necessary:*~~

Publication related data will be made available upon publication of the results.

Consolidated datasets will be made available at the end of the project.

Who will be able to access the data and under what conditions?

Indicate whether access to data will be open or restricted, for example to certain users or for certain purposes. Explain the procedures to control access (if applicable).

Tools to implement access conditions can be:

- licenses
- end-user agreements
- data sharing agreements

This License selector tool can help you find a suitable licence for data and software / code. Many data repositories offer a convenient selection of licences when you deposit data with them.

For data sharing agreements (advise and templates), contact KU Leuven Research & Development (Ird@kuleuven.be).

The datasets will be uploaded in Zenodo or similar as open access datasets under a Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license.

What are the expected costs for data sharing? How will the costs be covered?

Possible costs linked to open repositories and costs related to preparing data and uploading it will be covered by the project budget.

8. RESPONSIBILITIES

Who will be responsible for data documentation & metadata?

The applicant will be the responsible for data documentation & metadata.

Who will be responsible for data storage & back up during the project?

The applicant will be responsible for data storage & back up during the project, with support of KU Leuven central IT and local IT (SET-IT). The applicant has received information at the start of this project on the guidelines which apply in the hosting research group. For the implementation, the applicant can rely on the support of the Division LMSD's research manager Bert Pluymers.

Who will be responsible for ensuring data preservation and reuse ?

The applicant will be responsible for ensuring data preservation and reuse. Towards the end of the project, responsibility for long-term data preservation and reuse will be assigned in agreement with the KU Leuven Division LMSD research manager Bert Pluymers.

Who bears the end responsibility for updating & implementing this DMP?

This DMP is set up in agreement with the applicant's supervisor and co-supervisors. The applicant is responsible for updating and implementing this DMP.