
Condition monitoring of rotating machinery exploiting encoder and video measurements: development of novel encoder-based gear fault indicators by exploiting direct fault measurements as a ground truth

A Data Management Plan created using DMPonline.be

Creators: Toby Verwimp, n.n. n.n.

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Toby Verwimp

Grant number / URL: 1SE0123N

ID: 198650

Start date: 01-11-2022

End date: 31-10-2026

Project abstract:

Condition monitoring of rotating machinery gains importance in a number of industrial applications including wind turbines and automotive, in order to optimally schedule maintenance and to guarantee operation safety and production efficiency. Diagnostic indicators serve as tools to evaluate the condition of a rotating component and can be subdivided into two types. Indirect diagnostic indicators are based on signal processing and machine learning techniques applied on indirect measurements such as acceleration, sound or rotational speed. Direct diagnostic indicators can be defined as based on direct damage measurements by using images. Online monitoring systems exist using indirect diagnostic indicators, but these are rarely compared with direct damage measurements. Therefore, this proposal aims to develop a standalone direct monitoring (vision) system which is able to detect defects in rotating machinery at operational conditions, in order to serve as a ground truth for novel indirect monitoring techniques, with a focus on encoder measurements and pitting on gears. Firstly, a design for an online vision system is developed which should be robust against the harsh conditions in a gearbox complicating the visual access to the gear teeth. Secondly, novel encoder-based diagnostic indicators for gear monitoring are investigated and developed. The third step consists of integrating the two systems and methods on in-house test rigs with industrially relevant gearboxes.

Last modified: 27-04-2023

Condition monitoring of rotating machinery exploiting encoder and video measurements: development of novel encoder-based gear fault indicators by exploiting direct fault measurements as a ground truth

FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
UNSW Bearing run to failure	The run-to-failure experiments data were collected at the University of New South Wales in 2019-2020, regarding the development of bearing fault severity assessment methods.	Reuse existing data	Digital	Observational	.mat	<100GB	
Gear degradation data set	1. Images of gear teeth and rotary encoder signals for pitting detection 2. Damaged gear and tooth surface replicas	Generate new data	1. Digital 2. Physical	1. Observational and experimental	1. .ldsf .mat .bmp	1. <5TB	2. 1 gear and 13 replicas

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

UNSW Bearing run to failure:
Zhang, Hengcheng; Borghesani, Pietro; Zhuang, Siyuan; Peng, Zhongxiao (2021), "Bearing run-to-failure datasets of UNSW", Mendeley Data, V3, doi: 10.17632/h4df4mgrfb.3

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.

- Yes

Algorithms will be developed and implemented in order to detect damages in rotating machines. They will be tested and validated on above mentioned datasets and could have potential for commercial valorization.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

1. Code for simulations and code for data processing is accompanied by:

- a README.txt file explaining the code's purpose, inputs, outputs, connection to other files, assumptions, scope
- a documentation heading explaining the code's purpose, inputs, outputs, assumptions, paper on which it is based
- comments throughout the code for clarification

- units and a brief explanation after the declaration of each variable

2. Each newly generated data set is accompanied by an experiment report which describes/contains:

- the goal of the experiment
- the operational conditions
- the data acquisition settings and measurement procedures
- the structure of the generated data (folder structures and file names)
- the logbook

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

3. Data storage & back-up during the research project

Where will the data be stored?

- Both images and signals will be saved on an external hard drive.
- Both images and signals will be saved on a large volume storage: a shared file system available to members of the group via PCs and laptops managed by ICTS KU Leuven or our local KU Leuven IT department.
- Code will be saved on KU Leuven OneDrive and the research group's GitLab

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS.
- Back-up on external hard drive.

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

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The project will not use any personal data. All generated data is stored securely on KU Leuven's servers (GitLab, OneDrive) and behind proper authentication. During the project, the researcher and the promotor will evaluate how and when to share data. Sharing of data will occur through a secure channel such as Belnet Filesender.

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

€ 104,42 / TB / year (in blocks of 5 TB)

The costs will be covered by available internal funds.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data will be retained and possibly reused for future research projects.

Where will these data be archived (stored and curated for the long-term)?

The data will be stored on the external hard drives and on the large volume storages.

The data will possibly be uploaded to the KU Leuven repository Lirias and the division's GitLab.

Part of the data will possibly be published in dedicated data journals.

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

No costs for external hard drives (besides purchase cost).

Same cost for large volume storage (€ 104,42 / TB / year).

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, ...)

From the generated datasets, at least the raw data will be made available together with an experiment report.

If access is restricted, please specify who will be able to access the data and under what conditions.

Members of the research group will be able to access the data.

In future, the data can be shared in possible collaborations with other researchers/research groups.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

When the data will be shared, it will be through the Lirias repository.

When will the data be made available?

Some data will possibly be published together with a journal paper.

Which data usage licenses are you going to provide? If none, please explain why.

Probably, we will provide the Creative Commons Attribution (CC-BY-4.0) license.

This is the standard creative commons license that gives others maximum freedom to do what they want with our work (they are free to share and adapt), but they need to give appropriate credit and indicate if changes were made.

Do you intend to add a PID/DOL/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

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

No costs are expected for data sharing.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

The researcher will manage the code and datasets with proper metadata.

Who will manage data storage and backup during the research project?

The researcher will manage data storage and backup during the research project. This is supervised by Konstantinos Gryllias, the promotor.

Who will manage data preservation and sharing?

The researcher will manage data preservation and sharing during the project in collaboration with / supervised by Konstantinos Gryllias, the promotor. After the research project, the promotor will have this responsibility.

Who will update and implement this DMP?

The researcher will update and implement this DMP.