DTADD

Data Management Plan

Project Number: 101066739

Project Acronym: DTADD

Project Title: Digital Twin Anomaly Detection Decision-Making for Bridge Management

Experienced Researcher: Dr. Alejandro Jiménez Rios

Main Supervisor: Prof. Vagelis Plevris

Secondment Supervisor: Prof. Maria Nogal

April 2023

Table of Contents

List of tables

List of figures

No table of figures entries found.

1 Data Summary

The purpose of the data generated during the DTADD project is to assess the performance of different anomaly detection algorithms to detect damage or significant decay in existing bridges in Europe, and it is related to the achievement of research and innovation objective 1 of the DTADD fellowship.

The type of data generated will consist in a series of digital twin models which will be hosted at an open-source repository online (i.e., Github[1]). Those models will be used to run finite element simulations of idealized "*normal*" and "*not-normal*" bridge structural conditions. The output of those finite element methods simulations will be shared as .txt files in Zenodo[2].

At this point, it is not envisaged to re-use any existing data during the DTADD fellowship. All data in the DTADD project will be generated during its duration. Its origin will be the digital twin models generated and the results coming from the finite element method analysis performed.

The expected size of the generated data from the DTADD project would be less than 1 TB. This data might be useful for researchers working with digital twins and finite element method models as well as to those interested in developing and implementing state-of-the-art artificial intelligence algorithms.

2 FAIR Data

2.1 Making data findable, including provisions for metadata

All data produced during the DTADD fellowship will be assigned a DOI so that it can be easily findable. The naming conventions that will be used, will be established during the months when Task T1.1 of the Fellowship is performed (see proposal's Gantt chart). Furthermore, general keywords will be adopted during the DTADD fellowship to optimize the possibilities for re-use.

The tentative repositories that would be used for the storage of the data provide a versioning system that will aid to deliver clear version numbers for all the data generated during the DTADD fellowship. To the extent of the knowledge of the ER and supervisors, there is no metadata standard available for the disciplines of the DTADD project. Nevertheless, basic metadata will be assigned to all data generated in compliance with the FAIR Principles [1] and the general Dublin Core standard (Dublin Core | DCC) In case the scope of the project is expanded and any case study of a real heritage bridge is included, the MIDAS-Heritage standard will be adopted (MIDAS-Heritage | DCC).

2.2 Making data openly accessible

All data generated during the DTADD fellowship will be made openly available. The DTADD fellowship does not count with multi-beneficiaries, thus, there will be no need to keep any data closed. All data generated during the DTADD fellowship will be hosted at an open-source repository online, either GitHub for the digital twin models, or Zenodo, for the finite element models raw results files.

The data will be free to download using a common web browser. Once it has been downloaded, the user will require a licence of the software used during the generation of the digital twin models. On the other hand, the finite element models results, contained in raw .txt files, would be easily accessible with the use of any common text software. Moreover, additional

---

[1] https://github.com/
[2] https://zenodo.org/

documentation will be provided to indicate the necessary software to access the generated data during the DTADD project. Unfortunately, proprietary software is required for the creation of the digital twin models of the DTADD fellowship. But if any software or code snippets are generated for their creation and/or post-processing of the data, those resources would be made available as well (through a GitHub repository for example).

The experienced researcher (ER) has created several databases during past research projects, and he is also familiar with GitHub functionalities. Therefore, it is predicted that the selected repositories will comply with the databases requirements of the DTADD fellowship. Besides, by using those open repositories, there will be no restrictions in place, as far as the requirements of the DTADD project concern.

At this point of the fellowship, it has been deemed that the creation of a data access committee is not necessary. Unfortunately, no machine-readable license has been created at this point of the fellowship neither. This may change during future revisions of this DMP. Finally, the identity of the person accessing the data will be ascertained through the log-in mechanisms in place of the repositories that will be used by the DTADD fellowship.

2.3 Making data interoperable

Efforts will be done to ensure the interoperability of the generated data during the project. Unfortunately, due to the lack of a generally accepted standard in the discipline, its full operability cannot be guaranteed.

We will not follow any data of metadata vocabularies, standards, or methodologies to make our data interoperable. Furthermore, we will not be using standard vocabularies for all data types present in the data sets to allow interdisciplinary interoperability. If any commonly used ontology is found during the implementation of the DTADD fellowship, then mappings to it will be provided and made available

2.4 Increase data re-use (through clarifying licences)

All data generated during the DTADD fellowship will be licenced under the Open Data Commons Open Database License (ODbL) terms (Open Data Commons Open Database License (ODbL)). The data will be made available for re-use immediately after it is hosted in the mentioned repositories (GitHub and Zenodo).

These actions will allow for the data produced in the project to be usable by third parties. For more details about this, see the ODbL terms. Moreover, it is intended that the data remains re-usable forever. Finally, data quality will be ensured during the DTADD fellowship by applying bets practices on relevance, accuracy, consistency, timeliness, and compliance.

3 Allocation of Resources

No economic costs are involved for making the data of the DTADD fellowship FAIR, but a considerable time cost will be invested on it from the part of the ER. The ER will spend the necessary hours learning about FAIR principles and ensuring they are applied during the DTADD fellowship. Therefore, the ER will be the main responsible person for the data management of the project, but he will be closely supervised and helped by both supervisors of the fellowship as well.

The long-term preservation of the data generated during the DTADD has not been discussed yet at this point of the fellowship. This could be amended during a future revision of this DMP.

4 Data Security

All data will be available online by hosting it in the mentioned repositories. Furthermore, the ER will have a copy of all data in the laptop of the project which at the same time, will be hosted within his Microsoft OneDrive account, thus ensuring data is available at three different digital locations and could be recovered in case it is lost in one of them from either of the remaining two. For more details about the repositories' features regarding long-term data preservation and curation, visit their corresponding websites (GitHub, Zenodo).

5 Ethical Aspects

As identified in part B-2 of the DTADD proposal, no ethical issues are originally envisaged and in case they arise, they will be managed through a series of well-established guidelines and policies. No personal data will be collected during the DTADD fellowship.

6 Other Issues

No other national/funder/sectorial/departmental procedures for data management will be used.

References

1.   Wilkinson, M.D., et al., *The FAIR Guiding Principles for scientific data management and stewardship.* Scientific Data, 2016. **3**(1): p. 160018.

Annex A

Table 1. History of changes.

Version	Publication date	Change
1.0	April 2023	Initial version.