Probabilistic discovery of new microstructure-property relationships in advanced alloys

A Data Management Plan created using DMPonline

Creator: Joao Quinta da Fonseca

Affiliation: University of Manchester

Funder: Engineering and Physical Sciences Research Council (EPSRC)

Template: EPSRC Data Management Plan

ORCID iD: https://orcid.org/0000-0001-6063-8135

Project abstract:

Our aim is to revolutionize metallurgical practice, by enabling data-centred exploration and discovery of deformation mechanisms in advanced alloys by developing a novel, invertible, probabilistic modelling framework. Recently, there has been an explosion of the applications of machine learning methods to develop predictive models in material science. Yet often such predictive models are scientifically unsatisfactory since they are unidentifiable. That is, it is not possible, or at best extremely difficult, to identify the mechanism or physics which contribute to the observed response. This proposal flips the traditional machine learning approaches on their head, by building so called invertible neural networks for material discovery. Instead of building predictive models, maps from inputs to outputs, we will use high fidelity state-of-the-art-computational models and large, dense time-resolved data sets to train a specific probabilistic neural network architecture which is invertible. The resulting, probabilistic framework developed will be able to identify (with associated uncertainty) the inputs which directly lead to a given output. Therefore, the outputs from this proposal will develop a novel computational tool for the discovery of new microstructure-property relationships in advanced alloys

Last modified: 22-06-2020

Probabilistic discovery of new microstructure-property relationships in advanced alloys

UoM Project Details

Probabilistic discovery of new microstructure-property relationships in advanced alloys

- No

- Acquire new data
- Re-use existing data (please list below)

- Other repository or storage system (please provide details below)

Data will be stored on Zenodo

- > 8TB (please provide details)

Synchrotron data in particular will require a lot of space. We estimate 10-20 TB over the 5 years.

Data Collection

Experimental (physical data):
- text files
- binary files
- images
We will use standard, open file formats when possible.

Very varied techniques.

Documentation and Metadata

We are going to use at the beginning a simple set of metadata which will be the bare minimum needed by Zenodo data repository. We are going also to define a set of metadata specific to the project but following as much as possible metadata standard with eventually additional informations needed for the specificities of the project.

Ethics and Legal Compliance

There not ethical issue for this project.

Individual agreements with collabortators.

Storage and Backup

During data processing analysis the data will be stored in the Manchester RDS.
Once completed, data and analysis will be uploaded after creation and a first curation to be sure that they are consistent and with the proper metadata to the data repository (Zenodo). The upload will be eased by the usage of a dedicated software which will verify the presence of the needed metadata.

Most of the data will be place under open data license. For the one which will need to have an embargo or to be closed, Zenodo is providing the necessary tool to allows that. Sharing can be done specifically by asking the data creators through Zenodo.
Security is provided by Zenodo.

Selection and Preservation

Most of the data created by the project will be kept for future usage or to be use by other members of the collaboration. Zenodo is providing the space which allows that conservation.

Zenodo via the community facility

Data Sharing

Data will be shared between the members of the project as soon as they are produced. At the end of the project most of the data under embargo will be shared under an open data license. Some data, coming from industrial partners, could be keep closed but with a contact person if needed.

For some porojects with substantial industrial funding, yes.

Responsibilities and Resources

Joao Fonseca and Project RSE

RSE time and intermediate data storage (RDS).