Data publishing case study: from collection to publication

Engineering Nonlinearity

Version 1.6 March 2023



University of Bristol Research Data Service





INTRODUCTION

This document forms part of a series of data publishing case studies, each following a different project or research group in the Faculty of Engineering from initial data collection through to publication of findings. These case studies aim to showcase the ways that data sharing has been achieved across different projects with a variety of data formats and access restriction requirements.

Why publish data?

In both these examples, the authors are reporting work funded by EPSRC and are therefore expected to share their data as far as possible within ethical and/or commercial restraints. In addition, sharing data can increase the impact of a project, increase opportunities for collaboration, and thus maximise the benefit of public funding.

EPSRC guidance states that researchers should ensure that data underpinning published findings is available for scrutiny by others¹; however, what constitutes 'supporting data' is not defined. The following examples illustrate how researchers in the <u>Engineering</u> <u>Nonlinearity</u> project have interpreted this guidance.

EXAMPLE 1

Hill TL, Neild SA, Cammarano A. "An analytical approach for detecting isolated periodic solution branches in weakly-nonlinear structures". Journal of Sound and Vibration. 2016, 379:150-165. <u>doi:10.1016/j.jsv.2016.05.030</u>

Dataset DOI:

10.5523/bris.1nrlxw9m27k7g1s5pyuxomgs46

Data collection

In this study, oscillator backbone curves are modelled and the data collected is a series of files containing the coordinates representing numerically-derived solutions to the backbone curves. Working files are stored in FEN-RDS, the Faculty of Engineering Research Data Storage Facility.

Selection of data for publication

The authors have included plots visualising the modelled backbone curves as figures in the paper. For completeness they have decided to publish the data points underpinning these plots (i.e. the numericallyderived solutions), which in this case is the entire data collection associated with this work. These data were initially created in MATLAB, but for publication they have been exported to an open format text file as comma separated variables.

An argument could be made that since the equations defining the backbone curves are presented in the paper, along with the parameters used to derive the solutions, all the information required to verify the

¹ Clarifications of EPSRC expectations on research data management (UKRI) <u>https://www.ukri.org/about-us/epsrc/our-policies-and-standards/policy-</u> <u>framework-on-research-data/expectations/#contents-list</u>

plots is included in the paper already. By additionally publishing the underlying data points the authors have not only met but exceeded the minimum requirements for data sharing.

The data that have been selected for publication contain neither ethically nor commercially sensitive information, so are able to be openly shared.

Data deposit

In order to deposit in the University's data repository, data.bris, the lead researcher has applied for a personal allocation of storage space in the Research Data Storage Facility (deposit from FEN-RDS to data.bris is currently not possible). They have copied the files selected for publication into a pre-configured data-bris folder, added a readme.txt file and created a metadata record describing the deposit. Once the deposit has been checked by the Research Data Service, it has then been made public via data.bris (as no access restrictions were specified in the creation of the metadata record). A Digital Object Identifier (DOI) is provided so the published dataset can be easily cited.

Publication of findings

When writing up the study findings, the authors have included a data access statement in the acknowledgements section, which links to the deposited data and explains any access conditions (in this case, none):

Acknowledgements

The authors would like to acknowledge the support of the Engineering and Physical Scie nces Research Council. T.L.H. and S.A.N. are supported by EP/K005375/1. The data presented in this work are openly available from the University of Bristol repository at 10. 5523/bris.1nrlxw9m27k7g1s5pyuxomgs46.

Dataset usage

Within two years of publication, this dataset had received 64 unique views.

EXAMPLE 2

Shaw AD, Hill TL, Neild SA, Friswell MI. "Periodic responses of a structure with 3:1 internal resonance". Mechanical Systems and Signal Processing. 2016, 81:19-34. doi:10.1016/j.ymssp.2016.03.008

Dataset DOI: 10.5523/bris.1ovr9cp1whfbw1tfpql4hf10sq

Data collection

In this study, the nonlinear dynamics of a system are modelled and compared to experimental results generated by testing the response of that system to different inputs. The data collection in this case is (a) a series of files containing solutions for the modelled backbone curves (as in the previous example), and (b) a series of files containing experimental system input and output measures. Note that the experimental dataset (b) will contain a very large number of files as many repetitions will have been captured.

Again, working files have been stored in FEN-RDS, the Faculty of Engineering Research Data Storage Facility.

Selection of data for publication

The authors have decided that the experimental data (b) are of a higher priority for sharing than the modelled backbone curve data points (a). Furthermore, they have selected only the most representative examples from the large experimental dataset for publication (even this published subset is >1GB in size). These data were collected using custom software created in the LabVIEW development environment, but for publication purposes were converted to an open format text file as comma separated variables.

Again, the data selected for publication are neither ethically nor commercially sensitive, so no access restrictions are required.

Data deposit

As in the previous example, the data has been deposited in data.bris. However, to increase accessibility the authors have also chosen to deposit their data in Zenodo, a general-purpose data and publication repository. This is perfectly acceptable and is especially encouraged where there are subjectspecific repositories that may give greater visibility to a dataset.

Publication of findings

Again, the authors have cited the dataset DOI in the acknowledgements of the finished paper, allowing readers to easily locate the underlying data:

Acknowledgements The research leading to these results has received funding from EPSRC grant number EP/G036772/1. A.D. Shaw and M.I. Friswell received funding from the European Resear ch Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. [247045]. In addition, Simon Neild and Tom Hill are

supported by an EPSRC Fellowship (EP/K005375/1). The data presented in this work is openly available from the University of Bristol repository at http://dx.doi.org/10.5523/bris.1ovr9cp1whfbw1tfpql4hf10sq or from http://d x.doi.org/10.5281/zenodo.48133.

Dataset usage

Within two years of publication, this dataset received

119 unique views.

Acknowledgements

With thanks to Tom Hill for insights into study data collection and selection.