



# DISCIPLINARY DATA PUBLICATION GUIDES

ZOSIA BECKLES

UNIVERSITY OF BRISTOL

IDCC 2018

# BACKGROUND

What counts as **supporting data**?

Variation across disciplines in:

Data publication standards

Levels of detail in funder/publisher data policies

National/international guidelines understandably lack specificity

Guidance texts often too long for easy access

Leading to requests for support from researchers

# PROJECT AIMS

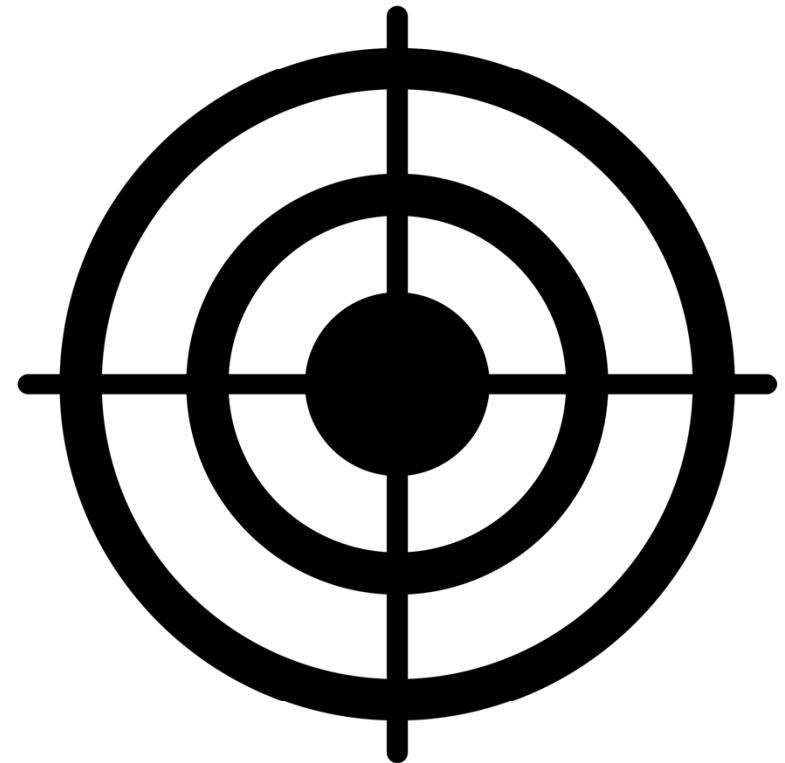
## Audience

Faculty of Engineering academics

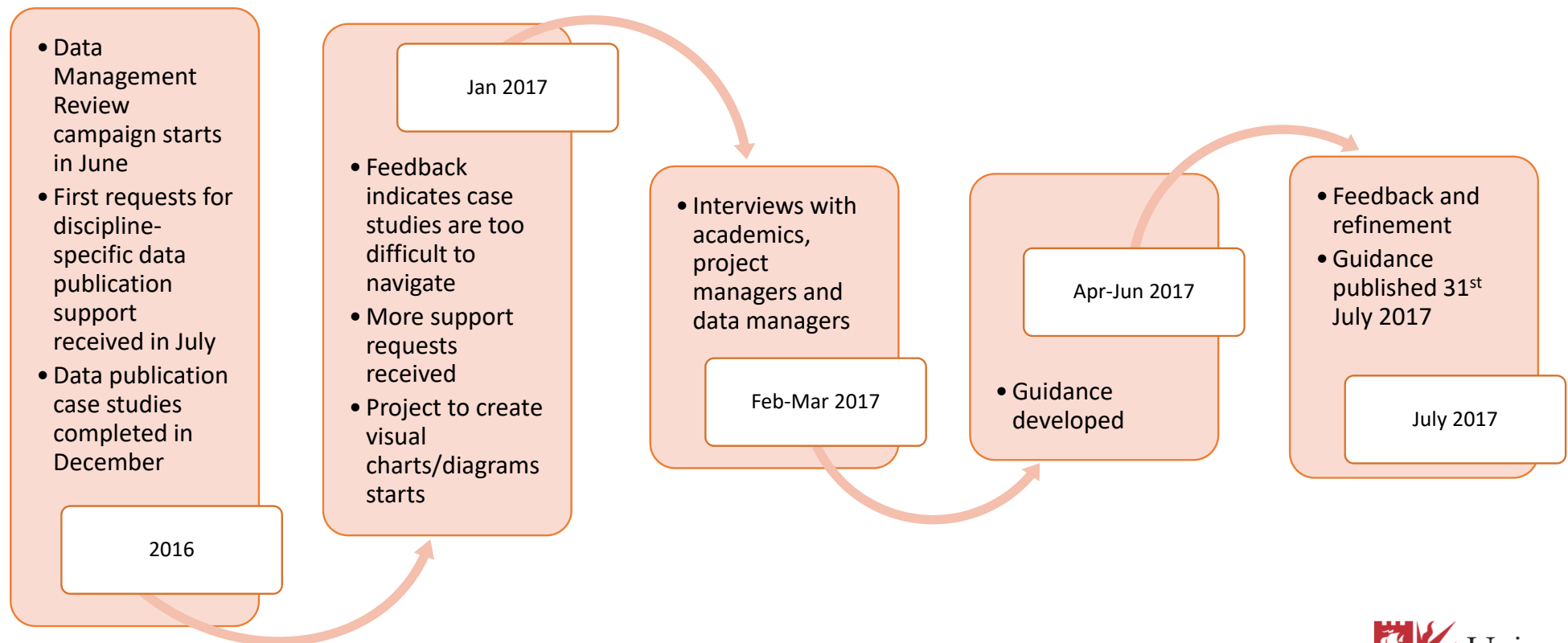
## Output

Short, visual guides covering

- Different types of paper
- Minimum data elements
- Sharing mechanisms



# METHODS



# DATA PUBLICATION ELEMENTS

## Paper type

- Modelling
- Experimental (physical)
- Review (including systematic review)
- Theory

## Data class

- Model set-up information
- Experiment set-up information
- Review set-up information
- Software or code
- Results
- Physical samples

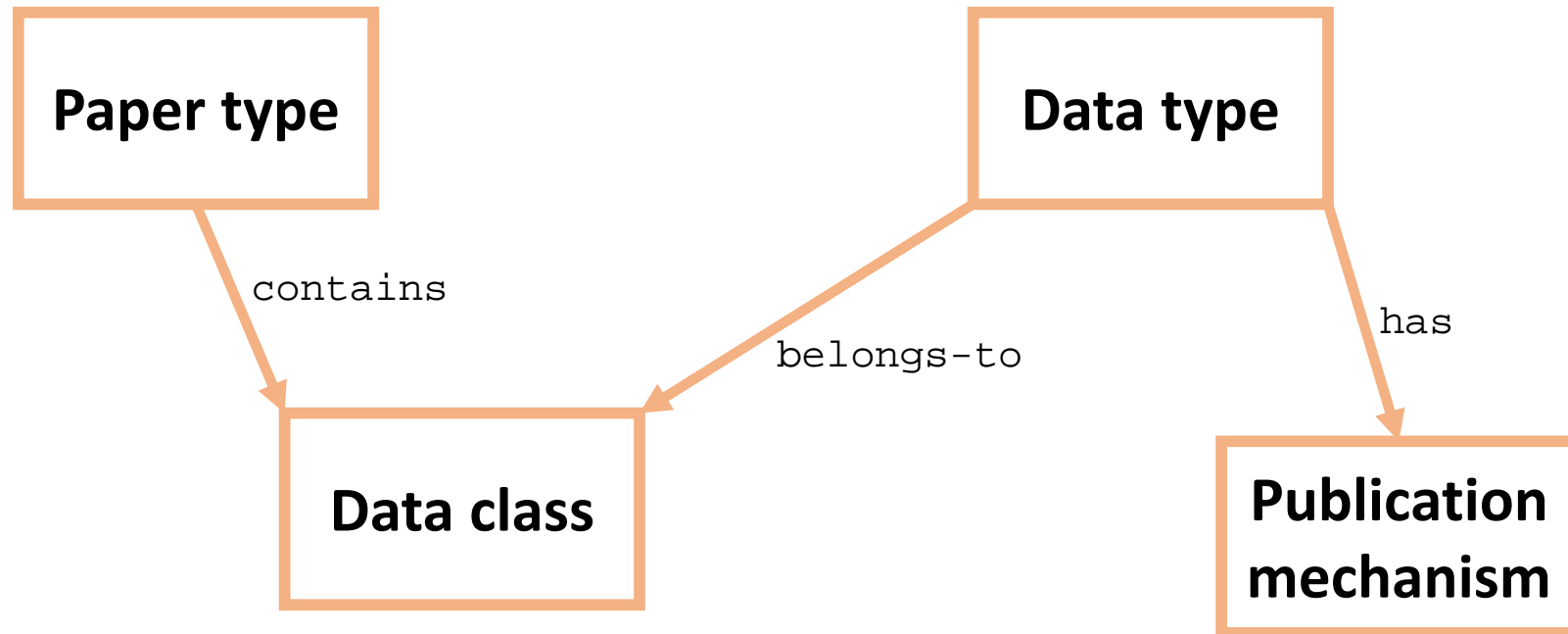
## Data type

- Third party software
- Code/software supporting workflow
- Code/software integral to study
- Model input or input conditions
- Description of model behaviour
- Complete model output
- Representative sample of model output
- Physical sample itself
- Description of physical sample preparation/capture method
- Experiment protocol
- Complete raw experimental data
- Representative sample of raw experimental data
- Complete processed data
- Representative sample of processed data
- Review protocol
- Summary statistics
- Derived data

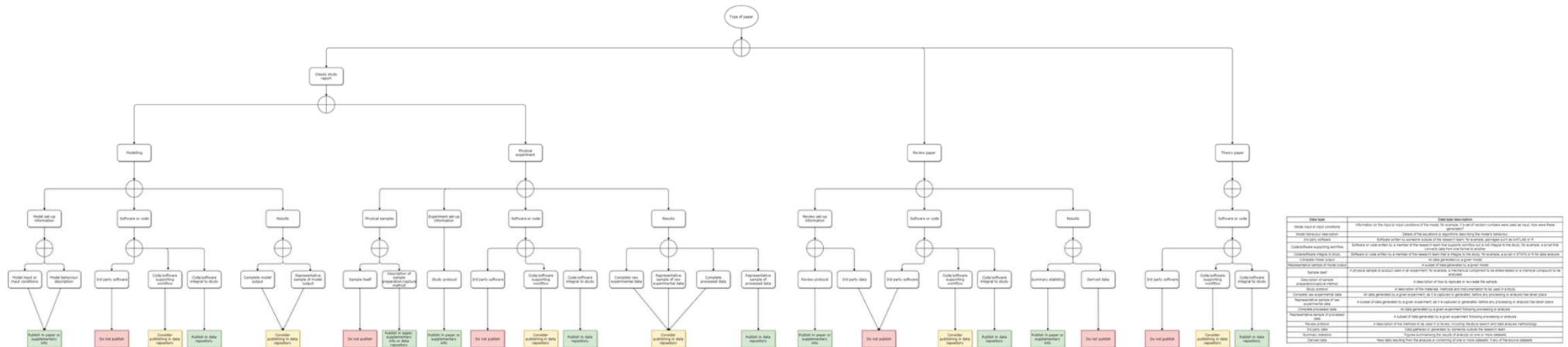
## Publication mechanism

- In data repository
- In supplementary information
- In paper
- Do not publish

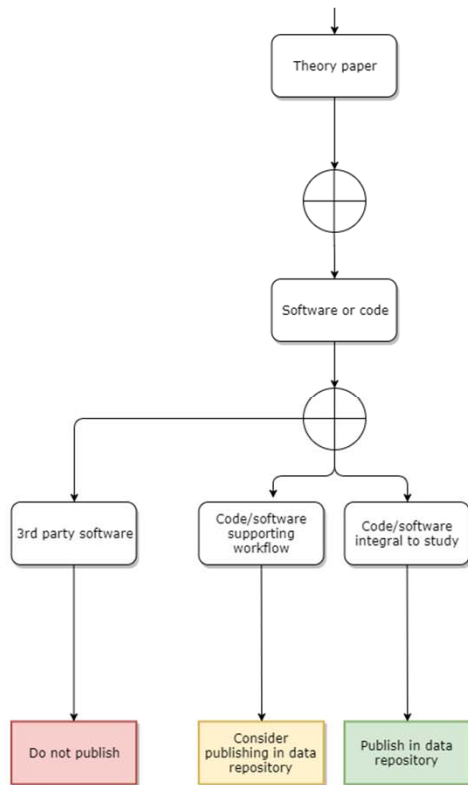
# ONTOLOGY



# DECISION TREE



# DECISION TREE DETAIL



Data type	Data type description
Model input or input conditions	Information on the input or input conditions of the model, for example, if a set of random numbers were used as input, how were these generated?
Model behaviour description	Details of the equations or algorithms describing the model's behaviour
3rd party software	Software written by someone outside of the research team; for example, packages such as MATLAB or R
Code/software supporting workflow	Software or code written by a member of the research team that supports workflow but is not integral to the study; for example, a script that converts data from one format to another
Code/software integral to study	Software or code written by a member of the research team that is integral to the study; for example, a script in STATA or R for data analysis
Complete model output	All data generated by a given model
Representative sample of model output	A subset of data generated by a given model
Sample itself	A physical sample or product used in an experiment; for example, a mechanical component to be stress-tested or a chemical compound to be analysed
Description of sample preparation/capture method	A description of how to replicate or re-create the sample
Study protocol	A description of the materials, methods and instrumentation to be used in a study
Complete raw experimental data	All data generated by a given experiment, as it is captured or generated, before any processing or analysis has taken place
Representative sample of raw experimental data	A subset of data generated by a given experiment, as it is captured or generated, before any processing or analysis has taken place
Complete processed data	All data generated by a given experiment following processing or analysis
Representative sample of processed data	A subset of data generated by a given experiment following processing or analysis
Review protocol	A description of the methods to be used in a review, including literature search and data analysis methodology
3rd party data	Data gathered or generated by someone outside the research team
Summary statistics	Figures summarising the results of analysis on one or more datasets
Derived data	New data resulting from the analysis or combining of one or more datasets. If any of the source datasets



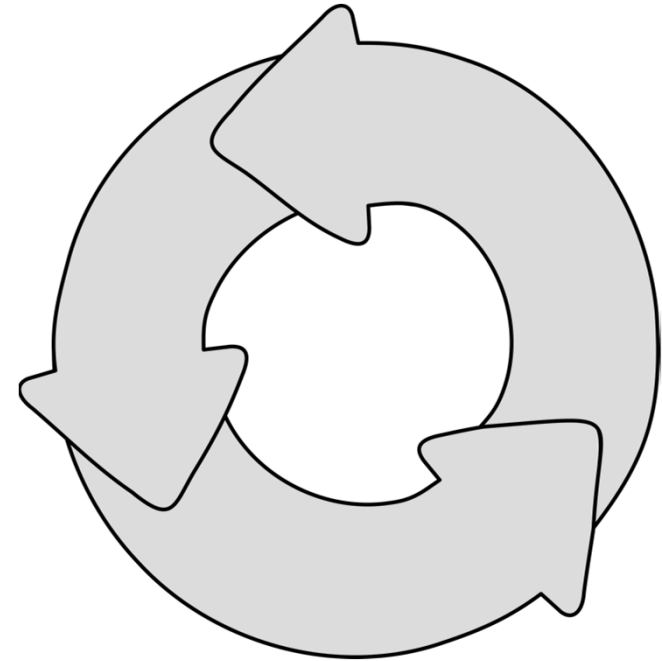
# ITERATIVE DEVELOPMENT

## Feedback sought from:

- Research group heads
- Research directors
- Project managers
- Research data management community

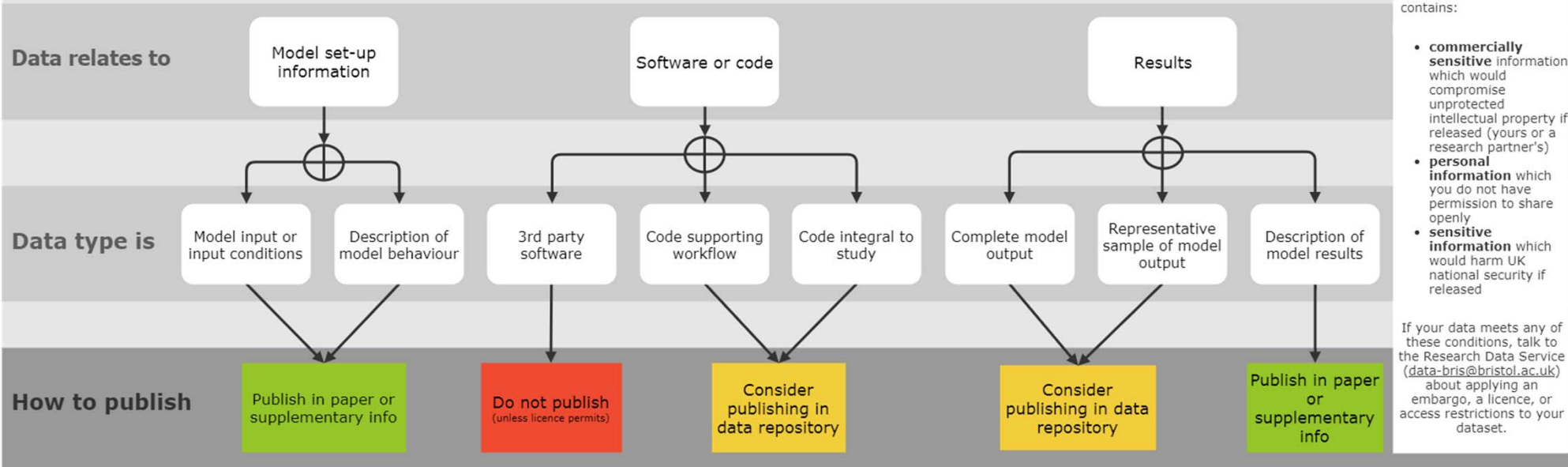
## Key changes:

- Theory paper type removed
- Information added on other restrictions to data sharing
- Decision tree split into separate diagrams



# FINAL DECISION TREES

## What data should I publish for a modelling study?



### Caution!

Before you publish a dataset, check whether it contains:

- **commercially sensitive** information which would compromise unprotected intellectual property if released (yours or a research partner's)
- **personal information** which you do not have permission to share openly
- **sensitive information** which would harm UK national security if released

If your data meets any of these conditions, talk to the Research Data Service ([data-bris@bristol.ac.uk](mailto:data-bris@bristol.ac.uk)) about applying an embargo, a licence, or access restrictions to your dataset.

# GUIDANCE RECEPTION AND FEEDBACK

Anecdotal and qualitative feedback generally favourable

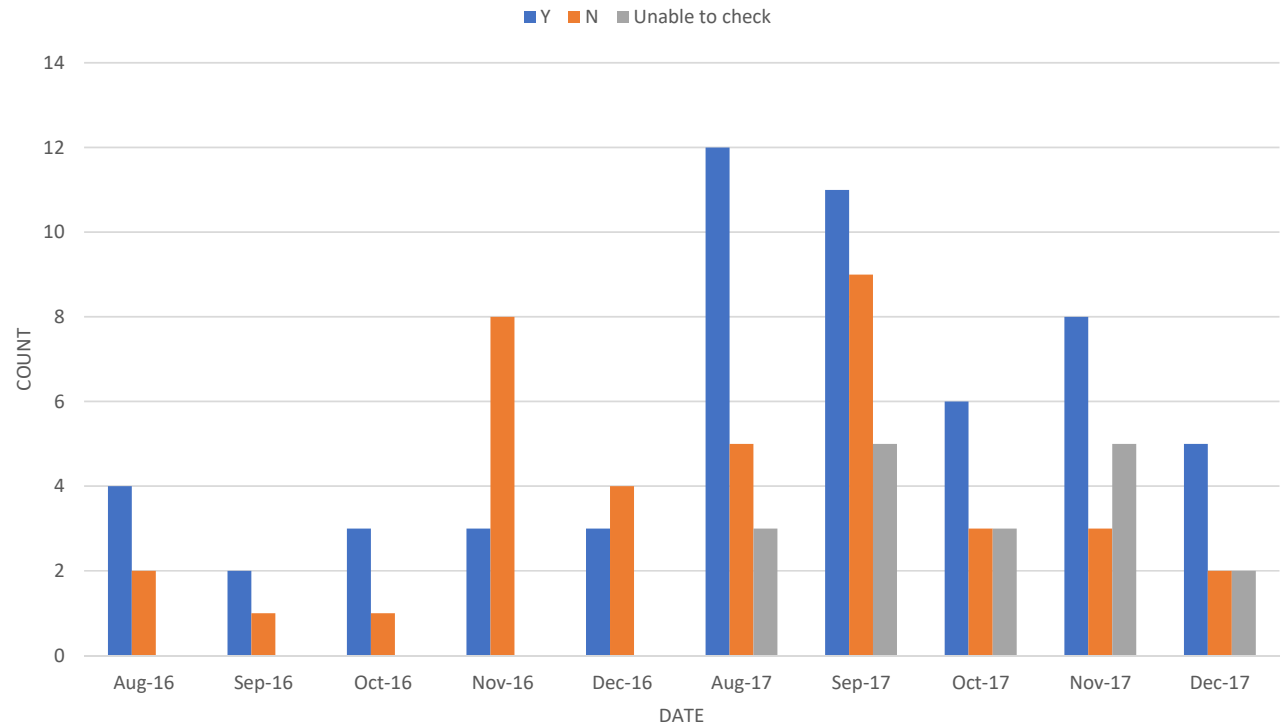
Effectiveness data hard to interpret

- Other training taking place simultaneously
- Short time since guidance released
- Generally positive effect on outcome measures

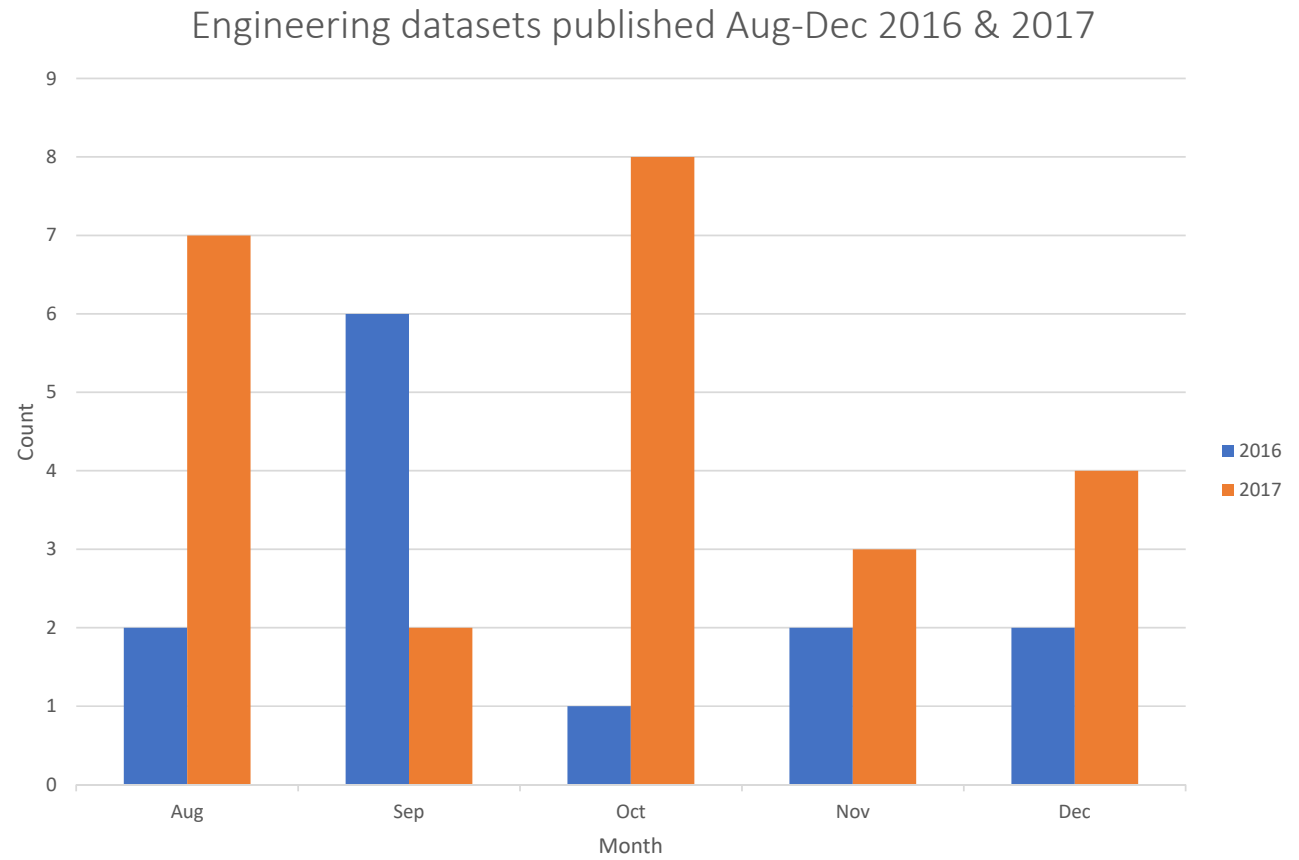
**Useful**  
**very useful reference**  
**practical overview**  
**very important**  
**self explanatory**  
**context**  
**good start**  
**useful**  
**providing clarity**

EFFECT OF  
GUIDANCE ON  
COMPLIANT  
DAS

Engineering Open Access papers with funder-compliant data access statements, 2016-2017



EFFECT OF  
GUIDANCE ON  
DATA  
PUBLICATIONS



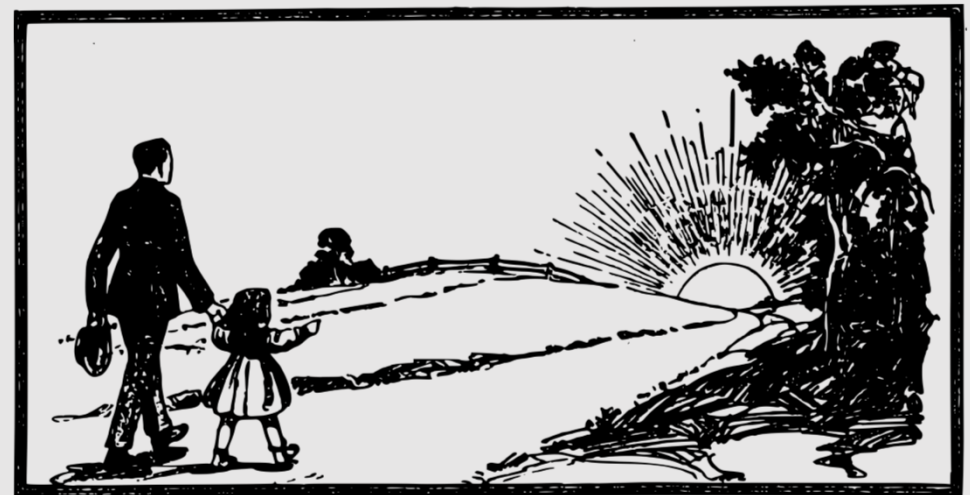
# KEY LEARNING POINTS AND NEXT STEPS

## Key issues

- Buy in from academics is critical
- Requires collection of a lot of background information
- Sustainability? Maintenance?

## Next steps

- Accessibility
- Other subjects
- Other data uses



# RESOURCES



Diagrams, contextual information, data management review template and effectiveness data are available at [10.6084/m9.figshare.5803266](https://10.6084/m9.figshare.5803266)

Diagrams were created using <https://www.draw.io/>

Please send comments and suggestions to [z.beckles@bristol.ac.uk](mailto:z.beckles@bristol.ac.uk)