Making a statement

Funder-compliant data access statements and support offered at the University of Bristol

Zosia Beckles,
Research Data Librarian

Stephen Gray,
Senior Research Data Librarian

Debra Hiom,
Assistant Director Research Services

Kirsty Merrett,
Assistant Research Data Librarian

Kellie Snow,
Research Data Librarian

Damian Steer,
Senior Technical Researcher

Research Data Service,
University of Bristol

data-bris@bristol.ac.uk

Abstract

There is an increased focus on the inclusion of data access statements in articles since the Engineering and Physical Sciences Research Council (EPSRC) expectation regarding this came into effect in May 2015. Data access statements are brief declarations in an article which cite any data underpinning claims and note any access restrictions placed upon these data, and the EPSRC expectation makes such a statement mandatory in any article acknowledging EPSRC funding. The University of Bristol has taken a variety of approaches to raise awareness of the requirements for data access statements, including promotional campaigns, email alerting and advocacy activities. In addition, we have developed guidance in several different formats. In addition to traditional discipline-specific guides, we have created an interactive web-based tool and a LaTeX package to help researchers create compliant statements. We have also worked collaboratively with other library teams, including Open Access, both to identify researchers in need of assistance and to monitor compliance. The latest figures indicate that 48% of researchers contacted have subsequently added compliant data access statements to their papers, and there has been a three-fold increase in the number of authors including compliant data access statements since our campaign to increase awareness began.

Background

The 2011 Research Councils UK (RCUK) Common Principles on Data Policy outlines high-level standards and expectations for the management and sharing of research data.[1] In particular, the policy includes a statement that 'published results should always include information on how to access the supporting data'. This is generally achieved by means of a data availability or data access statement: a brief declaration to readers citing any underpinning data, usually by means of a Digital Object Identifier (DOI), and also noting any access restrictions placed upon these data.

The RCUK principles are reflected in more or less detail in the individual data policies of the majority of UK funding bodies. Two bodies in particular have chosen to include a specific requirement for data access statements in their policies: from 1 May 2015, all articles acknowledging Engineering and Physical Sciences Research Council (EPSRC) funding are required to include a data access statement describing how any supporting data may be accessed, or alternatively, reasons why they cannot.[2] The Natural Environment Research Council (NERC) data policy includes a similar requirement for NERC-funded research.[3] The penalties for non-compliance are potentially very serious for both the researcher and the institution – the NERC data policy also stipulates that those who do not comply with the policy may have award payments withheld, or be deemed ineligible for future funding.

The University of Bristol Research Data Service (RDS) was established in 2013 to provide support in all aspects of research data management to academic staff and postgraduates at the university, including the management of the institutional research data repository, data.bris (https://data.bris.ac.uk/data/). We were therefore keen to ensure that all Bristol researchers were made aware of the specific data access statement requirements for EPSRC and NERC-funded research, and to offer support and advice where needed.

We initially focused on broadly promoting EPSRC data management expectations, including producing data access statements. As EPSRC-funded researchers are not required to submit a data management plan (DMP) with their grant applications, it was felt that there was a greater risk of essential data management tasks being overlooked in this group. Once the EPSRC campaign was under way, we extended aspects of it to NERC-funded researchers as well. We used a variety of strategies to ensure that our message was widely received, including working with other teams in the university library and across the university. These methods are discussed below.

Making a statement
Funder-compliant data access statements and support offered at the University of Bristol
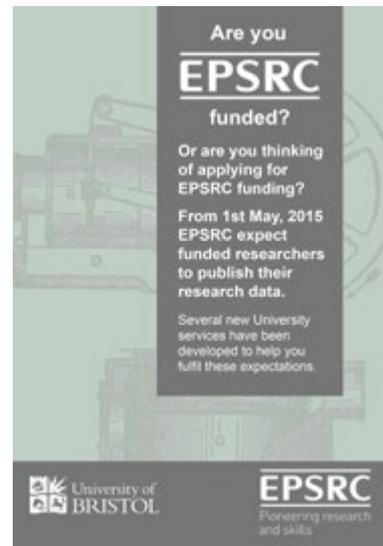
Awareness-raising strategies

We launched a multi-phase promotional campaign in early 2015; this focused primarily on the faculty of engineering, which had been identified as the faculty in receipt of most EPSRC funding. Firstly, posters, banners and digital signage were placed in key locations in the faculty buildings, outlining the four main tasks required to meet EPSRC expectations on data management:

- preservation of data for at least ten years post-project
- inclusion of a data access statement in all published papers
- inclusion of data management costs in grant applications
- creation of a data management plan (EPSRC require this to be completed, but not submitted as part of a grant application.)

The poster can be seen in Fig. 1. However, we were aware that there were still likely to be some members of staff who would not see these, so we also created a series of leaflets which were sent to key contacts and disseminated at School assemblies and workshops (Fig. 2).

EPSRC data management poster

EPSRC leaflet 1

EPSRC postcard – reverse

In May 2016, a year after the EPSRC expectations came into effect, we launched a second series of posters and leaflets with greater focus on the requirements for data access statements. The leaflets were sent out to research staff with their payslips and included examples of compliant data access statements on the reverse side, in order to help researchers construct their own statements (Fig. 3).

In late 2016 we also extended the poster campaign to selected non-engineering research staff. By consulting EPSRC's Grants on the Web portal (http://gow.epsrc.ac.uk/) we were able to identify individuals and research groups in receipt of significant EPSRC funding, and sent copies of our poster to them.

The other aspect of our awareness strategy involved collaborative working with the library's Open Access (OA) team. Researchers wishing to claim article-processing charges from the university's RCUK and Charity Open Access Fund block grants apply to the OA team. From April 2015, the OA team sent a standard email to researchers acknowledging EPSRC or NERC funding in their claim form to remind them of the data access statement requirements, and forwarded details of these applications to the Research Data Service.

Finally, in 2015, a one-off series of workshops was added to the regular training schedule specifically addressing the EPSRC expectations on research data and what researchers would need to do to fulfil them. We followed this up by updating our usual workshops and other advocacy and outreach activities (aimed at researchers from all faculties) to include explicit mention of the requirements for data access statements, including examples of different types of statement.

Guidance and support

We created a number of web-based guidance documents and tools to support our awareness-raising activities. These were intended to act as stand-alone sources of advice that we could direct researchers to outside our training workshops.

Firstly, we brought together more examples of compliant data access statements for major subject specialisms. Following the discipline divisions at Bristol, we produced three brief guides to data access statements in arts and social sciences, biomedicine and health science, and science and engineering. We used examples from publications by Bristol researchers in almost all instances to ensure that they were relevant to our audience. The examples covered a variety of situations, including:

- papers without data
- review (secondary analysis) papers
- data as supplementary material
- data published in the institutional repository
- data published in an external repository
- data published with access restrictions
- data described in a separate data paper

We received feedback that the data access statement examples were helpful, but that more guidance on how to decide which data to publish was a higher priority. After consultation with senior academics in the faculty of engineering we created more in-depth case studies to describe this aspect in more detail. The case studies were initially intended to track a single project through the whole research lifecycle, but the selected research groups expressed caution about publicising their data management plans. We therefore limited the case studies to the later stages of the research lifecycle: data collection, selection of data for publication, data deposit, publication of findings and re-use.

We identified research groups in the faculty of engineering who had published data in the institutional data repository (http://data.bris.ac.uk/data), and from these chose datasets representing a variety of sub-disciplines and access restrictions. For example, the Sensor Platform for Healthcare in a Residential Environment (SPHERE) case study describes the publication of a dataset containing information on human participants which was published as a restricted dataset available only to bona fide researchers in line with the limits set in the participant consent forms.

We interviewed the researchers in each team responsible for collecting the data and writing up the findings in order to find out how they had decided

what data to publish. This process highlighted the intensely subjective nature of determining what constitutes 'supporting data', as it became apparent that even for papers reporting work from the same project, different types of data were selected for publication.

This subjectivity is demonstrated in our second case study, based on two papers published by the Engineering Nonlinearity group. This group had taken a nuanced, case-by-case approach to data selection; in the first example, a numerical modelling paper, they had published files containing the data points underpinning each plot in the paper. The second paper involved numerically modelling a system and comparing this to experimental results generated by testing the response of that system to different inputs. In this instance, the group decided that although the plot data points were available, it would be more relevant to readers of the second paper to publish the most representative experimental results.

Since completion of these two case studies, other engineering groups have expressed interest in having case studies based on their own data types. Lastly, we produced a range of tools to automatically generate compliant data access statements. For those who are very new to data access statements, we have developed an interactive tool using Squiffy (http://textadventures.co.uk/squiffy). This asks a series of questions around data ownership and sensitivity to guide researchers to the appropriate statement. The data access statement generator tool will be incorporated into an online research data guidance tool being developed jointly with Bath, Exeter and Cardiff that will be available for user testing in February 2017. More information on this tool will be presented at the 12th International Digital Curation Conference (http://www.dcc.ac.uk/drupal/events/idcc17).

For those who were more confident in deciding what kind of statement they needed, but merely wanted help with wording, we developed a LaTeX package that allowed researchers to generate a compliant statement at any point in their document by entering a simple command. This is available both as a ShareLaTeX template and as a zipped package for download.
All guidance documents and LaTeX templates are available on our website at https://data.bris.ac.uk/sharingdata/.

Monitoring

Measuring impact is often difficult, but we were able to work with the OA team again to gain insight into the rates of compliance. They have monitored the compliance of published Gold papers with RCUK's policy since 2013, and as part of this were collecting information on data access statements in papers acknowledging RCUK funding. They shared this information with us, and we were able to match this against researchers to whom we had sent reminder emails (part two of our awareness campaign) in order to gauge effectiveness.

From these data, it can be seen that our awareness campaign and guidance and support activities are correlated with a significant increase in the proportion of papers with compliant data access statements. So far, in 2016 48% of researchers who were sent reminders subsequently added compliant data access statements to their papers, and we have noticed that some 'frequent flyers' are now adding statements without reminders. From 2013 to 2015 there was a gradual increase in the proportion of papers with compliant data access statements (7.4 % to 10.8%) across all faculties; however, from 2015 to 2016, when the additional support was put into place, this figure increased three-fold, to 34.8% (Fig. 4). When only the publications acknowledging funding from EPSRC and NERC are considered, the increase from 2015-2016 is even more dramatic: from 9.2% to 37.0% for EPSRC, and from 18.9% to 60.1% for NERC (Fig. 5).
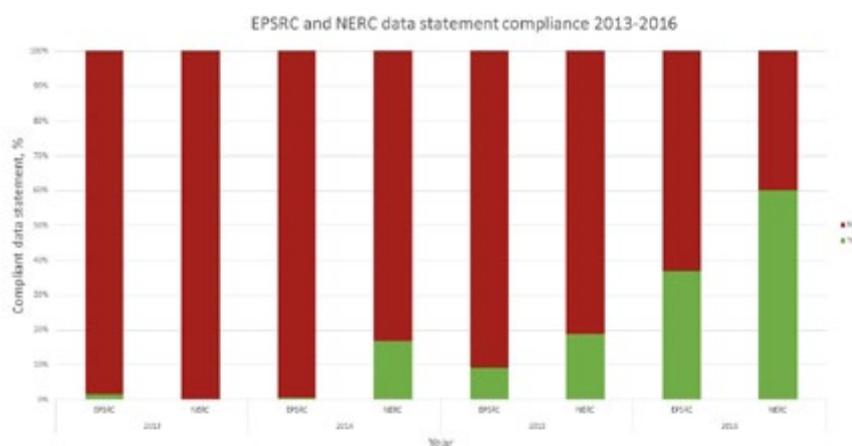
Making a statement
Funder-compliant data access statements and support offered at the University of Bristol

Overall data statement compliance 2013-2016

EPSRC and NERC data statement compliance 2013-2016

Next steps

Whilst there have been significant improvements in the targeted groups, overall more than half of eligible papers are still non-compliant, so continued reminders and encouragement are necessary. Much of the guidance that has been produced so far has been weighted towards meeting EPSRC requirements, and it would be useful to expand the work to cover RCUK-wide data access statement stipulations. Finally, it is essential that we continue working closely with other library teams and faculties across the university in order to ensure that we provide relevant, timely advice to researchers.

References

1   RCUK, 2011. RCUK Common Principles on Data Policy [online]. Available from: http://www.rcuk.ac.uk/research/datapolicy/ [accessed 7 December 2016]

2   EPSRC, 2014. Clarifications of EPSRC expectations on research data management [online]. Available from: https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement/ [accessed 1 December 2016]

3   NERC, 2010. NERC Data Policy [online]. Available from: http://www.nerc.ac.uk/research/sites/data/policy/data-policy/ [accessed 1 December 2016]