IJDC | *Brief Report*

ESRC Research Data Policy in a Changing Landscape

Paul Allanson

Angela Daly

Awais Elahi

Alistair Geddes

Maeve Malone

Niamh Nic Daeid

Lucille Tetley-Brown

University of Dundee

Abstract

The Economic and Social Research Council (ESRC) is the UK's largest funder of economic, social, behavioural, and human data science. ESRC research data policy is 'intended to support ESRC grant holders who collect, produce and re-use data by defining researchers' roles and responsibilities, as well as the roles and responsibilities of the ESRC and its data service providers.' This paper reports on an independent review of this policy, commissioned and funded by the ESRC, and carried out by the authors, to make recommendations for an updated policy in the light of changes to both the data and legal/policy landscape. Following an initial scoping review, the study comprised an online survey of stakeholder views that was followed up by a series of focus groups, and an analysis of a sample of data management plans (DMPs). In this brief report, we concentrate on those aspects of our review, both in process and in substance, which are of most relevance to the data curation and data management community, and outline the next steps in the policy review process.

Introduction

The ESRC is the UK's largest funder of economic, social, behavioural, and human data science, providing support both for original research and for the management of data collections and the provision of data services. ESRC research data policy is 'intended to support ESRC grant holders who collect, produce and re-use data by defining researchers' roles and responsibilities, as well as the roles and responsibilities of the ESRC and its data service providers' (ESRC, n.d.), and was most recently revised in 2018. A key element of the policy since its introduction in the 1990s has been the default requirement for new data created through ESRC-funded research to be deposited with the UK Data Service (UKDS) or another suitable depository once the ESRC grant ends (Van den Eynden, & Corti, 2017). However, much has changed since 2018 with regards to the kinds of research data produced by social science researchers – going far beyond conventional survey data for which the policy was originally designed. The contemporary data landscape that the policy must now span consists of a growing array of more diverse forms of data, both primary and secondary, as well as software code and algorithms. In addition, there have been significant changes in key aspects of the legal and policy landscape pertaining to research data – such as the UK implementation of the European Union (EU) General Data Protection Regulation in 2018, and the exit of the UK from the EU in 2020.

This paper reports on an independent review of ESRC research data policy that was commissioned and funded by the ESRC (grant no: ES/Y005546/1). The goal of this review was to identify and evaluate changes to the data, legal and policy landscape as regards their interaction with and impact on ESRC research data policy, and to recommend modifications to the policy and associated practice to ensure its fitness-for-purpose. The evidence-based review was carried out by the authors (Allanson et al., 2024a, 2024b) who constituted a multidisciplinary team from the University of Dundee led by the socio-legal scholar Angela Daly, between August 2023 and April 2024.

In this brief report, we concentrate on aspects of our review which are of most relevance to the data curation and management community, including our engagement with this community and the findings from our review having the greatest impact on this community. We first provide an overall summary of our review before focusing on those aspects most relevant for data curation and management, and finish by outlining next steps in the policy review process.

Independent Review of ESRC Research Data Policy

Our review was designed as a multi-phase study, beginning with an initial scoping of the broader context of research data practices, policies, and legislation that informed subsequent engagement activities with UK-resident stakeholders in the ESRC research data community. These activities comprised an online survey, a series of focus groups, and a set of ESRC-facilitated discussions (under the Chatham House rule) with representatives from data-focussed investments receiving significant ESRC funding. The survey asked for respondents' views about the current design and implementation of ESRC research data policy and the need for change due to the growing importance of new and emerging forms of data, with these issues explored further in the focus groups. In addition to these engagement activities, we evaluated a sample of research data management plans (DMPs) prepared as part of ESRC research funding applications and provided to us by the ESRC and the University of Dundee on a confidential basis. The purpose of an individual DMP is to document how research data will be managed both during and after a funded research project, with our evaluation being particularly timely given changes to the DMP requirement in the ESRC funding application process following the Tickell Review (Tickell, 2022). The bulk of this work took place between September and December 2023.

Our stakeholder engagement was aimed at including a wide cross-section of people and roles in the scope of influence of ESRC research data policy. Social science researchers producing and/or using data formed a major grouping in this cross-section. Those involved in data curation and management activities, or supporting such activities were also included in all three main engagement activities. The findings support a view that the existing ESRC research data policy is generally seen as helpful, with the default policy position on requiring data deposit among the positively regarded aspects, and with DMPs considered to be a useful tool. However, evidence gathered from the work also confirmed a case for the policy and its implementation to better recognise the contemporary diversity of social science data associated with different methods, approaches and disciplines. Another area identified as in need of improvement relates to uncertainty around how the research community can best produce open research data, connected to a perceived under-acknowledgement in the existing policy of funded data which cannot be made fully open for reasons such as privacy and confidentiality. Furthermore, there was a strong view that funding support was insufficient for the work involved in preparing data adequately for sharing and re-use.

Detailed recommendations to support the development and implementation of an updated ESRC research data policy are grouped under four themes in our final report (Allanson et al., 2024a). First, we recommend a set of aims and guiding principles for an updated ESRC research data policy, to enhance its coverage of roles, responsibilities, and rights of all relevant parties, as well as its forward applicability, at least over near-term timescales. These aims and principles are founded on the view that the ESRC should continue to aim to maximise the use of research data that it funds for public benefit, and that it should also exercise a key role in building public trust in research data use by promoting those benefits. Second, we set out several measures that the ESRC should take to ensure the achievement of these goals through the terms and conditions that it sets for funded researchers and major research investments, and through the guidance and advice it provides these and its other constituents. Third, we advocate maximum coherence among the UK Research and Innovation (UKRI) funding councils, of which the ESRC is one, as regards research data, leveraging synergies and reducing gaps and differences where possible. Specifically, we recommend that a move towards policies for distinct types of research data or contexts is considered by UKRI, rather than council-specific research data policy documents - where they exist - as is currently the case. Finally, we make a case for regular policy review, to ensure that the policy remains fit for purpose, and propose several topics warranting further research to strengthen longer-term policy development and updates.

Focus on Data Curation and Management

We have engaged with the data curation and management community in the UK throughout the lifetime of our review and are grateful to all those who participated in our stakeholder engagement activities. Preliminary findings from the review were presented as work-in-progress at the 18th International Data Curation Conference (IDCC24) in Edinburgh in February 2024, with a more detailed presentation of DMP-specific aspects of the work at the DMPOnline conference on 21 May 2024. Here we focus on the recommendations in the final report that are directly relevant to the curation and management of research data.

Recommendation #2 on enhancing data deposit proposes that the ESRC should develop additional guidance on documenting data, and that it should also consider means to mandate use of persistent dataset identifiers, to help mainstream and normalise good data citation practices. This responds to stakeholder requests for more guidance on data documentation and to ensure more consistent and best practice use of persistent identifiers across social science research data in the UK as currently there is some unevenness in both knowledge and practice. Recommendation #7 on enhancing DMPs sets out the view that DMPs are retained as part of the ESRC grant application process, but they should be reworked, to enhance their effectiveness. This recognises stakeholder feedback affirming the importance and usefulness of DMPs for researcher and institutional planning, while acknowledging previous practice can be improved. Recommendation #10 is that the ESRC make more funding available to cover the

costs of data management, which should not come from organisational overheads but should be a separate, funded budget item, reflecting stakeholder feedback that the real costs of data management are higher than currently budgeted, especially in the context of pressure on university finances and resources. Recommendation #16 on alignment focusses on exploring options for the increased standardisation and harmonisation of metadata by ESRC and its investments, and the production of a common set of data-related terminology, taking into consideration developments on both these fronts by other key players, such as the Medical Research Council (another UKRI council) and the UK Office for National Statistics.

In our report we expand upon the DMP recommendation, providing further detail on what reworked DMPs should look like. First, evidence from our review suggests that the planned 500 word limit on DMPs in the ESRC funding application process should be increased quite substantially – we suggest a new limit of 1,500 words – to enable all relevant information to be captured, yet without making DMP writing overly burdensome, Second, we recommend the implementation of a living DMP process in which the original plan is amended as necessary over the entire research data lifecycle to ensure it remains up to date, with examples provided of 'trigger points' for reviewing and updating the plan. DMP preparation and management processes should be facilitated by developing interoperability between the Funding Service (through which ESRC funding applications are currently made) and online DMP tools such as the Digitial Curation Centre's DMPOnline, and by making DMPs machine actionable. Finally, we propose that DMPs should include consideration of a new 'risk of harm' principle concerning the sharing of riskier and/or more harmful data that may need to be subject to more restrictions on access and re-use, as well as recommending extended guidance to address knowledge gaps on certain matters, including intellectual property, data protection and international collaborations.

Next steps

The final report was presented at a hybrid launch event held on 15 May 2024 with over 100 attendees. The event was attended in person by ESRC representatives including the Deputy Chief Executive and included a panel session discussing issues raised by our review findings and recommendations. The ESRC is currently working on a response to our review, including inter alia how it will address our recommendations and their implementation, and is aiming to move onto the next stage of actioning the recommendations in September/October 2024.

Acknowledgements

References

Allanson, P., Daly, A., Elahi, A., Geddes, A., Malone, M., Nic Daeid, N., & Tetley-Brown, L. (2024a). *Doing ESRC Data Sharing Better: Final Report*. (Science, Policy and Law series; No. 2). Leverhulme Research Centre for Forensic Science, University of Dundee. doi: 10.20933/100001312

Allanson, P., Daly, A., Elahi, A., Geddes, A., Malone, M., Nic Daeid, N., & Tetley-Brown, L. (2024b). *Doing ESRC Data Sharing Better: Appendices*. (Science, Policy and Law series; No. 2). Leverhulme Research Centre for Forensic Science, University of Dundee. doi: 10.15132/10000253

Economic and Social Research Council (n.d.) *ESRC Research Data Policy.* Retrieved from https://www.ukri.org/wp-content/uploads/2021/07/ESRC-200721-ResearchDataPolicy.pdf

Tickell, A. (2022). *Independent Review of Research Bureaucracy - Final Report.* Retrieved from https://assets.publishing.service.gov.uk/media/62e234da8fa8f5033275fc32/independent-review-research-bureaucracy-final-report.pdf

Van den Eynden, V., & Corti, L. (2017). Advancing research data publishing practices for the social sciences: from archive activity to empowering researchers. *International Journal on Digital Libraries, 18*, 113–121. doi: 10.1007/s00799-016-0177-3