Responsible Data Governance in Projects: Applying a Responsible Research and Innovation (RRI) Framework

Tilimbe Jiya*1

Abstract: There are large amounts of data that are collected, processed and exchanged in projects and they form the basis of innovation and fact-finding activities that are at the core of project activities. Nevertheless, such data can be irresponsibly used, managed and governed during its lifecycle, starting from creation all the way to its use. Given the importance of data and the risk of being misused, it is important to understand how responsibilities for data management can or should be designed in projects. This paper showcases how the concept of Responsible Research and Innovation (RRI) can be implemented in projects to mitigate irresponsible data management and use. The paper draws on the practices undertaken in two EU funded research projects to help practitioners find ways of dealing with data management related issues.

Keywords: responsible research and innovation; responsible data governance; data management; data management; project management

Submitted: January 28th, 2021 / Approved: May 23rd, 2021

Introduction

Data is a crucial aspect for projects in modern societies. Data acts as the fuel for project activities as they are the basis of research and knowledge, providing insights with a potential to improve our welfare, create new businesses, allow societies to develop new networks and lead richer lives. At the same time, data can be misused, contain biases, constitute the basis of misinformation and rumours, or stifle freedom. To ensure that projects in modern societies can reap the benefits of ever-increasing amounts of data and resultant insights while keeping the disadvantages in check, responsibilities around the creation and use of data must be defined.

This paper explores the question of how responsibilities for data governance can or should be designed. It does this by drawing on the concept of responsible research and innovation (RRI), a term widely used in research governance discussions (Aicardi et al. 2018). The paper argues that RRI offers a suitable basis to think about responsible data use, at least in a research context. After introducing RRI as the theoretical backbone, the paper outlines the current key issues and concerns regarding data. It then offers an account of two European Union (EU) funded research projects that employ ideas from RRI in dealing with data-related questions. By contrasting these two examples, the paper identifies current gaps in responsibility for data use in projects.

The paper makes a theoretical contribution by introducing the theory and current discussions within RRI to the discourse concerning the moral impact of data. By providing an account of current practices from the two projects, it furthermore helps project practitioners to find ways of dealing with data-related concerns.

Responsibility and Responsible Research and Innovation

This section provides a brief introduction to RRI, followed by a discussion of challenges and concerns that are typically linked to data. The term ‘responsibility’ points to a willingness and ability to answer. One is typically responsible for something or somebody (Stahl 2004), which means that one is held to account for what happens to that something or somebody. There are many different types of responsibility, such as role responsibility, legal responsibility, parental responsibility and many others. What most of them have in common is a moral connotation. Being responsible is a moral quality; in turn, being irresponsible is (morally) bad.

But what does this mean in practice? How can the generally positive concept of responsibility translate into clear actions that ensure that data is generated and used in a desirable and morally acceptable way? The answer to this question consists of a reference to RRI, though there is no space in this brief paper to detail RRI extensively. Briefly, it is a term widely used by research funders and policymakers (European Commission 2012), to ensure that processes and products of research and innovation are desirable, acceptable and sustainable (von Schomberg, European Commission., and Directorate-General for Research and Innovation. 2011). It is important to note that this broad concept has found its way into the discussion of the social consequences of information and communication technology (Jirotka et al. 2017) and thus a link to data use is already established.

One important aspect of RRI is its ambition to ensure that research and innovation promote the public good. One way of achieving this is by pursuing the UN’s Sustainable Development Goals (UN 2018). These and other grand challenges thus guide RRI. In practice, RRI offers more guidance on what to do and how to do it. There are
different views on what exactly constitutes RRI. However, for this paper, the focus is on the view that is promoted by the UK Engineering and Physical Sciences Research Council (EPSRC). The view promoted by the EPSRC emphasises a set of activities summarised by the acronym AREA (Owen 2014). This stands for anticipation, reflection, engagement and action. A consideration of the AREA activities in research and innovation has the potential to address some of the concerns about data use.

Concerns about data use

Data in projects is being generated at an ever-increasing volume and faster rate than ever before. Projects are being implemented in what people call a data-rich society, whereby data is generated by our conscious actions (e.g. making calls to clients using your mobile phone to gather requirements) and background or automated systems (e.g. communication technologies ping a telecoms tower to establish and keep a signal ready for when you want to communicate with internal and external project stakeholders).

There are concerns over the generation, collection and manipulation of data that is used in project. Further, there are worries over how project members are looking after the data that is shared during its lifecycle including the requirements gathering to project closure phases (Cellan-Jones Rory, 2018). Other concerns relate to the implications of using data that is corrupted or collected in such a way that it can be seen as ‘wrong’ (Martin, Borah, and Palmatier, 2017). Also, there are questions about the consequences of biases that are introduced when collecting and manipulating data in projects (August, 2020). These concerns reflect some key issues that relate to data governance in projects.

Key issues with data use

Accountability

Issues around accountability for data and its management lie at the heart of responsible data governance. In the context of ICT research projects such as those used in this paper, questions of responsibility and data exist within a tension across ethical imperatives: related calls for open data, open access, and open science to drive innovation and foster transparency (Salerno et al. 2017), and the need for data security, concerns around data protection and privacy, and requirements for compliance with legislation such as the General Data Protection Regulation (GDPR); in the case of Europe, but certainly other parts of the world have similar regulations. Data security is a significant concern with regard to data management and accountability (Siddiqua et al. 2016).

In addition to negotiating this tension, data management and governance processes throughout the data lifecycle must also account for issues of intellectual property and credit for contributors. “Ownership” of data is not always straightforward, and may well be shared or contested. For personal data in a European context, it is also necessary to have a legal basis for processing (e.g. informed consent). These needs can partly be addressed through ’meta-responsibility’ (Rainey et al. 2017), and the creation of distributed governance frameworks and bodies responsible for managing compliance, data policy, protocols, and workflows which incorporate the tenets of RRI (Aicardi et al. 2018).

Transparency

The project management community and, perhaps more importantly, the wider public, must be assured that the ways in which the data they provide is handled are appropriate and that those projects exploiting the data can be held accountable for its use.

Recently, there has been an upsurge of big data use in projects and the algorithms which fuel them, have become embedded in everyday project management activities, from using internet search engines, social media feeds, to predictive analysis in managing projects using big data analytics. Complexity breeds opacity, and as the algorithms become more impenetrable it becomes increasingly difficult to identify the individual data which inform decisions that are made in projects (Richards and King 2013).

The aim should be that big data analytics used in projects is as transparent as possible, but what is possible is up for debate. Analytics at the scale of big data is a complex assemblage of code, mathematics, statistical models and, of course, a certain amount of human intervention. What is possible and what is desirable may not necessarily be completely congruent.

Human rights

The advancement of data use in developing technologies such as Smart Information Systems (SIS) that are used in projects is currently generating a mix of excitement and concern across the globe. These are emerging technologies use AI, Big Data and predictive analytics. For instance, there have been ongoing debates amongst stakeholders over the threats of data use in SIS, while optimists have argued that the use of data in SIS might be directed towards solving societal challenges (EC 2018) and therefore improving activities and flow of actions within projects. However, these narratives have the potential to distract from the idea that the use of data in many of these already-widespread technologies can have distinct implications for human rights.

The intersections between human rights and data-driven technologies have been growing. Perhaps the most noticeable have been expectations of a significant decrease in employment due to automation. Also, there have been concerns over the privacy impacts of SIS that use data to predict behaviour and recognise patterns, for example, with facial recognition software. The manipulation of data in SIS pose risks of discrimination through exacerbation of bias and the effects of some predictive decision-making methods that affect society today such as policing, criminal justice systems and access to essential economic and social services (Zwitter 2018). An example, can be seen from the Gender Shades project which looked at how leading technology companies’ commercial AI systems significantly mis-gender women and darker skinned individuals as a result of the priorities, preferences, and prejudices of those who have the power to mould artificial intelligence (Buolamwini and Gebru 2018).
In addition, there is a concern on how data use is affecting the way projects in our society should be managed without violating human rights. The use of data, particularly in project SIS, is leading towards a state where people are becoming incapacitated.

**Policy and regulation**

Currently, the question of how responsibilities for data should be designed is high on the agenda of policy-makers. Recent policy documents focusing on 'data revolution' outline a range of ideas for new institutions and regulations to ensure responsible data use. A novel idea of creating Data Trusts (Hall and Pesenti 2017) – proven and trusted frameworks and agreements – is being developed in the UK as a part of its Industrial Strategy that defines Artificial Intelligence and Data Economy as one of its four Grand Challenges (HM Government 2017).

The establishment of Data Trusts is envisaged as a way to improve trust and ease around sharing data between organisations and projects as well as to ensure that data exchanges are secure and mutually beneficial. These Trusts should allow data holders and users to collaborate more easily by providing a repeatable framework for sharing data in a fair, safe and equitable way. According to the UK government, ‘these frameworks will ensure that all parties involved have defined rights and responsibilities towards the data and individuals’ data, and other sensitive data is protected’ (HM Government 2018).

**Method and data sources**

This paper draws insights from work in two cases, namely the Human Brain Project (HBP) and Shaping the ethical dimensions of smart information systems – a European perspective (SHERPA) project. These insights are used to discuss how RRI can be put into practice in dealing with data-related questions. In addition to data obtained from the empirical insights, the paper is also informed by documents produced by the projects thus far to understand how RRI can be used as a framework to mitigate the negative aspects of responsibility issues for data use in projects.

**Cases**

To answer the research question of how responsibilities for data can or should be designed in projects, the paper applies the AREA framework to the analysis of two EU-funded research projects as real-life cases for understanding how RRI can be used to address data use in project management. Two cases provide a set of empirical insights to judge the value of RRI in mitigating responsibility issues in data use in projects. The background to each of these cases is briefly described below:

SHERPA is a 42-month, EU-funded initiative which began in May 2018, and involves 11 partners across Europe. The project will investigate, analyse and synthesise people’s understanding of how smart information systems (SIS) impact ethics and human rights issues. SIS is the combination of artificial intelligence and big data analytics. The project will develop novel ways of understanding and addressing SIS challenges, evaluate these with stakeholders, and advocate for the most desirable and sustainable solutions (SHERPA 2018).

The HBP is building a research infrastructure to help advance neuroscience, medicine and computing. The HBP is one of the two largest scientific projects ever funded by the European Union. The 10-year Project began in 2013 and directly employs more than 800 people in over 100 universities, teaching hospitals and research centres in 20 countries. The project’s infrastructure includes six ICT research Platforms: Neuroinformatics, Brain Simulation, High-Performance Analytics and Computing, Medical Informatics, Neuromorphic Computing and Neuorobotics. To address major societal implications, the HBP also includes the Ethics and Society division.

These cases are compared and contrasted to identify similarities and differences in data use.

**Data sources**

In both cases, data sources included researcher’s experience and documents. The author, a researcher involved in the two research projects examined these data to assess whether and to what extent RRI is being used to deal with responsibility issues for data use in each case. Also, documents that are part of the deliverables for the two cases were used to complement these understandings and satisfy the aim of this paper. These documents included public documents, reports and internal project documents.

**Data governance responsibilities in practice**

**SHERPA Project**

In an era where research innovation and ethics go hand in hand, the challenge remains to steer innovations in such a way that technology delivers its benefits without compromising people’s integrity, rights and privacy.

The SHERPA project aims to investigate and analyse our understanding of how smart information systems (such as AI and big data analytics) impact ethics and human rights. Many of the recognised issues of SIS are linked to the use of personal data, hence protecting such data is an important step to avoid negative consequences, and the GDPR will also help to deal with this. The SHERPA project adopts RRI to understand and address the ethical issues of SIS. RRI is a suitable approach that has been shown to be relevant for key issues such as privacy (B. C. Stahl 2013; Bernd Stahl and Yaghmaei 2015). More importantly, it provides a lens that allows for a broad understanding of the responsibility issues that arise from technologies such as SIS and how these could be addressed. With regard to data use in the SHERPA project, RRI will ensure that those different stakeholders collectively reflect on potential consequences of data use and deliberate on solutions that could be used to mitigate the consequences. Some of these consequences are related to trust, security and privacy.

As part of the collective reflection and action, SHERPA will engage a range of expert stakeholders (see Table 1). Part of the stakeholder engagement will involve consultation on the alignment of SIS with relevant standards such as those developed through CEN/CENELEC and ISO/IEEE initiatives. CEN and CENELEC are international non-profit associations and are officially recognised as European
Standardisation Organisations. Their national members and committees work jointly to develop and define standards, and support the implementation of European legislation. It is envisaged that the findings from SHERPA will help to contribute to existing initiatives or start new standardisation efforts.

The consultation will identify areas where SHERPA would contribute to existing initiatives and frameworks, or start new standardisation efforts. In addition to this, the project will devise case studies and scenarios, to identify concerns and preferred solutions. It is one of SHERPA’s aims to raise awareness of these issues, including those that relate to data use or governance through European consensus. Therefore, the results will be used to present technical and regulatory options, in order to advocate, promote and implement the most promising solutions through targeted dissemination activities. One such example will be to assess the potential of standardising good practices or quality criteria, to ensure human rights and ethics considerations are at the forefront when handling or processing human data.

Human Brain Project
Much of the data handled by HBP staff (whether generated by activity within the project, imported into or else exported outside of the project) is sensitive and can cover several modalities. Instantiated within the HBP are structures and processes meant to govern the usage of data within the project. Considering the duration, operational and scientific breadth of the project, these structures must be compositionally robust, methodologically proactive and reactive in relation to the potential for cultural, societal and political change throughout the lifetime of the project.

The HBP sought to engage with the challenges posed by data in the Data Protection and Privacy Opinion (HBP 2018b) by the Ethics and Society subproject of the HBP as part of wider RRI-related efforts in the subproject (see Table 2). This document (published in 2016) makes a number of suggestions with regard to data protection within the project.

One suggestion was that the project should seek to set up a Data Governance Committee to review privacy and data protection processes within the project; therefore, the HBP set up the Data Governance Working Group (DGWG). The DGWG is a collection of representatives from all twelve sub-projects who support data governance and management infrastructure development. Outputs from the DGWG include the forthcoming Data Policy Manual and the Data Policy Quick Guide, the latter synthesises and highlights relevant policies for HBP scientists while the former contains much of the detail and background to these policies. The work of the DGWG, and the guidance within both the Data Policy Quick Guide (HBP 2018a) and the forthcoming Data Policy Manual help to inform and guide HBP scientists working to generate, import or export data within the project.

The HBP also employs a Data Protection Officer, who provides expert input into data governance issues including Data Protection Impact Assessments, and plays a key contributory role in the DGWG. These structures help to foster a scientific community that is conscious of data governance challenges across the project and recognises the importance of responsible data handling in the HBP.

Synthesis
Having examined the two cases individually, the AREA framework (Aicardi et al. 2018) is used to compare how the two projects seek to anticipate, reflect, engage and act in ensuring responsible research, respectively. To that effect Table 1 is presented below to indicate how the AREA framework is applied to both projects; SHERPA project and the HBP.

Table 1. Applying the AREA framework to SHERPA and HBP

| AREA framework | Application to the SHERPA Project | Application to the Human Brain Project |
|----------------|---------------------------------|---------------------------------------|
| Anticipation   | ✅ Collaboratively finding appropriate ways to effectively reduce the risk of data misuse that could result in tensions between SIS and the protection of human rights. | ✅ The Foresight Lab produced a series of reports on Future Computing and Robotics, Future Medicine, and Future Neuroscience (Aicardi et al. 2018), in addition to developing scenarios. |
|                | ✅ Findings to generate new insights into how to responsibly research, develop and use of data in SIS. | ✅ Consultations with a broad range of stakeholders to inform the actions of the Ethics and Society Subproject and the HBP to shape the development of ethics governance structures in the HBP (Rainey et al. 2017). |
| Reflection     | ✅ Production of ethical guidelines, and protocols for research and innovation in SIS. | ✅ The Ethics Support group and the Ethics Rapporteurs work with the independent Ethics Advisory board to incorporate reflection on ethical issues into project-wide protocols and processes. |
|                | ✅ Examination of some regulatory models (e.g. government, co-regulation, self-regulation or polycentric regulation) that are relevant to artificial intelligence and big data. | ✅ Conceptual reflection is also supported by the activities of the Neuroethics and Philosophy group, who collaborate with neuroscientists to enhance the social relevance of HBP research. |
|                | ✅ Assessment of the potential of standardising good practices or quality criteria to ensure ethics and human rights in SIS. | |

ISSN: 0718-2724. (http://jotmi.org)
Journal of Technology Management & Innovation © Universidad Alberto Hurtado, Facultad de Economía y Negocios.
## Conclusion

As data becomes increasingly integral to every facet of the project’s lifecycle, it becomes ever more important that those individuals and organisations that utilise data in projects are ready to engage with the complexity of the attribution of responsibility for data use at every point in its lifecycle. If data-use in projects goes unchecked and without associated governance structures, there will be implications for human rights; for instance, consider the opacity of decision-making processes using big data, and the impingement on privacy through the utilisation of project data, to name a few examples. These are issues that those individuals and institutions which utilise data in their projects should consider: not only potential future problems, but also current challenges faced by those in the wider public from whom the data is gathered, and who will be ultimately affected by the outcomes of their activity.

This paper has argued that data governance encourages responsible data use in projects. It has proposed that RRI provides a conceptual framework to deal with concerns that relate to the responsible governance of data use. This has been showcased by applying the AREA framework to understand how data governance could be used and is currently being used in research by looking at two EU research projects at different stages but exhibiting similar issues surrounding data use. Despite differences, the projects have things in common. Both projects operationalise the idea that data governance is not just rules, policy documents and standards but also the involvement of different groups, roles, communities, discourse and exchanges. Such involvement ensures that there is distributed or shared responsibility in data use within projects. Looking at one dimension of the AREA framework provides an illustration: both projects hold consultations with experts, which can help inform the instantiation of data governance structures.

The paper contributes to the practice of data governance in projects by providing an applicable and relevant way of reflecting on how responsible data use could be integrated into the management of projects. Using RRI as an approach for responsible governance is important if the concept and its related practices are to be relevant beyond research projects such as those used as cases and applied to projects in other contexts.

**Funding:** This paper has been informed by the research in the SHERPA project which has received funding from the European Union’s Horizon 2020 research and innovation programme, under grant agreement No 786641 and the Human Brain Project which has also received funding from the same programme, under grant agreements No 720270 (SGA1) and 785907 (SGA2).

**Acknowledgements:** The author would like to acknowledge the contribution of all SHERPA and Human Brain projects participants and all their activities to the ideas that underpin this paper.

## References

Aicardi, C., Fothergill, B. T., Rainey, S., Stahl, B. C., & Harris, E. (2018). Accompanying technology development in the Human Brain Project: From foresight to ethics management. *Futures*. https://doi.org/10.1016/j.futures.2018.01.005

August, T., Fox, R., Roy, D. B., & Pocock, M. J. (2020). Data-derived metrics describing the behaviour of field-based citizen scientists provide insights for project design and modelling bias. *Scientific reports*, 10(1), 1-12.

Buolamwini, J., & Gebru, T. (2018). *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*. In *Proceedings of Machine Learning Research* (Vol. 81, pp. 1–15).
Cellan-Jones, D. L., Rory. (2018, June 27). Ticketmaster admits hack attack. *BBC News*. Retrieved from https://www.bbc.com/news/technology-44628874

Cellan-Jones, Z. K., Dave Lee, Rory. (2018, March 21). Cambridge Analytica: The story so far. *BBC News*. Retrieved from https://www.bbc.com/news/technology-43465968

EC. (2018). Communication From The Commission To The European Parliament, The European Council, The Council, The European Economic And Social Committee And The Committee Of The Regions Artificial Intelligence For Europe, *Brussels (Com(2018) No. 237 Final)*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN

European Commission. (2012). Responsible Research and Innovation. Europe’s ability to respond to societal challenges. *European Commission*. Retrieved from https://ec.europa.eu/research/swafs/pdf/pub_public_engagement/responsible-research-and-innovation-leaflet_en.pdf

Hall, W., & Pesenti, J. (2017). Growing the Artificial Intelligence Industry in the UK. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf

HBP. (2018a). HBP Data Policy Quick Guide. Retrieved from https://sos-ch-dk-2.exo.io/public-website-production/filer_public/25/2a/252ac9d0-d408-41fc-8fae-bab89dae5ee6/hbp_data_policy_quick_guide.pdf

HBP. (2018b). Human Brain Project Ethics and Society Opinion: Data Protection and Privacy f. Retrieved from https://sos-ch-dk-2.exo.io/public-website-production/filer_public/42/e2/42e28dca-6d5d-4513-9771-88ab71fc3ce1/data_protection.pdf

HM Government. (2017). *Industrial Strategy: Building a Britain fit for the future.* (p. 256).

HM Government. (2018). *Industrial Strategy Artificial Intelligence Sector Deal*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/702810/180425_BEIS_AI_Sector_Deal___4_.pdf

Jirotka, M., Grimpe, B., Stahl, B., Eden, G., & Hartwood, M. (2017). Responsible research and innovation in the digital age. *Communications of the ACM*, 60(5), 62–68.

Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data Privacy: Effects on Customer and Firm Performance. *Journal of Marketing*, 81(1), 36–58. https://doi.org/10.1509/jm.15.0497

Owen, R. (2014). The UK Engineering and Physical Sciences Research Council’s commitment to a framework for responsible innovation. *Journal of Responsible Innovation*, 1(1), 113–117. https://doi.org/10.1080/23299460.2014.882065

Rainey, S., Stahl, B. C., Shaw, M. C., & Reinsborough, M. (2017). Ethics Management and Responsible Research and Innovation in the Human Brain Project. Edward Elgar Publishing. Retrieved from https://www.dora.dmu.ac.uk/xmlui/handle/2086/13145

Richards, N., & King, J. (2013, September 3). Three Paradoxes of Big Data. Retrieved July 27, 2018, from https://www.stanfordlawreview.org/online/privacy-and-big-data-three-paradoxes-of-big-data/

Salerno, J., Knoppers, B. M., Lee, L. M., Hlaing, W. M., & Goodman, K. W. (2017). Ethics, big data and computing in epidemiology and public health. *Annals of Epidemiology*, 27(5), 297–301. https://doi.org/10.1016/j.annepidem.2017.05.002

SHERPA. (2018). SHERPA - Understanding and Analysing Smart Information Systems. Retrieved July 27, 2018, from https://www.project-sherpa.eu/

Siddiq, A., Hashem, I. A. T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., & Nasaruddin, F. (2016). A survey of big data management: Taxonomy and state-of-the-art. *Journal of Network and Computer Applications*, 71, 151–166. https://doi.org/10.1016/j.jnca.2016.04.008

Stahl, B. (2004). Responsible Management of Information Systems. Hershey, PA: IGI Publishing.

Stahl, B. C. (2013). Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy*, 40(6), 708–716. https://doi.org/10.1093/scipol/sci067

Stahl, B. C., Rainey, S., Harris, E., & Fothergill, B. T. (2018). The role of ethics in data governance of large neuro-ICT projects. *Journal of the American Medical Informatics Association*: JAMIA. https://doi.org/10.1093/jamia/ocy040

Stahl, Bernd, & Yaghmaei, E. (2015). The Role of Privacy in the Framework for Responsible Research and Innovation in ICT for Health, Demographic Change and Ageing (Vol. AICT-498). Springer International Publishing. https://doi.org/10.1007/978-3-319-55783-0_8

UN. (2018). Agenda for Sustainable Development. Retrieved July 27, 2018, from https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

Stahl, B. (2012). Responsible Research and Innovation. Europe’s ability to respond to societal challenges. *European Commission*. Retrieved from https://ec.europa.eu/research/swafs/pdf/pub_public_engagement/responsible-research-and-innovation-leaflet_en.pdf

Richards, N., & King, J. (2013, September 3). Three Paradoxes of Big Data. Retrieved July 27, 2018, from https://www.stanfordlawreview.org/online/privacy-and-big-data-three-paradoxes-of-big-data/
