Supporting Responsible Research and Innovation within a University-based digital research programme: reflections from the "hoRRIzon" project

Virginia Portillo , Peter Craigon , Dowthwaite Liz , Chris Greenhalgh , Elvira Pérez-Vallejos

PII: S2666-6596(22)00022-1
DOI: https://doi.org/10.1016/j.jrt.2022.100045
Reference: JRT 100045

To appear in: *Journal of Responsible Technology*

Received date: 7 January 2022
Revised date: 23 May 2022
Accepted date: 23 August 2022

Please cite this article as: Virginia Portillo , Peter Craigon , Dowthwaite Liz , Chris Greenhalgh , Elvira Pérez-Vallejos , Supporting Responsible Research and Innovation within a University-based digital research programme: reflections from the "hoRRIzon" project, *Journal of Responsible Technology* (2022), doi: https://doi.org/10.1016/j.jrt.2022.100045

Supporting Responsible Research and Innovation within a University-based digital research programme: reflections from the "hoRRIzon" project

Virginia Portillo[a†], Peter Craigon[a,b], Liz Dowthwaite[a], Chris Greenhalgh[a,c], Elvira Pérez-Vallejos[a,d]

[a] Horizon Digital Economy Research Institute, University of Nottingham, Nottingham, UK
[b] Future Food Beacon of Excellence, University of Nottingham, Nottingham, UK
[c] School of Computer Science, University of Nottingham, Nottingham, UK
[d] Faculty of Medicine and Health Sciences, University of Nottingham, Nottingham, UK

[†]corresponding author: Virginia.Portillo@nottingham.ac.uk. Horizon Digital Economy Research Institute, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU

Co-authors' emails:
Peter.Craigon4@nottingham.ac.uk;
Liz.Dowthwaite@nottingham.ac.uk;
Chris.Greenhalgh@nottingham.ac.uk;
Elvira.Perez@nottingham.ac.uk

Abstract

Integration of Responsible Research and Innovation (RRI) principles into a research project is key to ensure outputs are ethically acceptable and socially desirable. However, translating RRI principles into practice is challenging as there are no recipes that fit all projects. Drawing on interviews and case studies with researchers, external partners and support staff associated with two linked programmes of research addressing trustworthy autonomous systems and the ethical use of personal data, this reflective paper provides a bottom-up perspective on RRI as it is understood and "done" within the focussed, often relatively short, research projects that largely comprise the research programmes. Existing RRI frameworks and tools were found to provide useful resources for use in supported RRI activities. However, effective use is dependent on facilitation by experienced staff, and time pressure and timing within a project can limit the application of RRI. Research institutions and funding bodies could provide more practical and sustained support for RRI practice – throughout the research lifespan - including incentives for RRI and clearer RRI/RI roles and career pathways, helping to develop an embedded community of RRI practice throughout the digital research community and beyond.

Supporting Responsible Research and Innovation within a University-based digital research programme: reflections from the [XXX] project

Highlights

- Time and timing are critical elements when applying RRI in relatively short research projects
- The value of RRI tools is strongly enhanced by facilitation by research staff with expertise in RRI
- Embedding a culture of RRI demands practical and sustained support and resources from institutions and funding bodies

Abstract

Integration of Responsible Research and Innovation (RRI) principles into a research project is key to ensure outputs are ethically acceptable and socially desirable. However, translating RRI principles into practice is challenging as there are no recipes that fit all projects. Drawing on interviews and case studies with researchers, external partners and support staff associated with two linked programmes of research addressing trustworthy autonomous systems and the ethical use of personal data, this reflective paper provides a bottom-up perspective on RRI as it is understood and "done" within the focussed, often relatively short, research projects that largely comprise the research programmes. Existing RRI frameworks and tools were found to provide useful resources for use in supported RRI activities. However, effective use is dependent on facilitation by experienced staff, and time pressure and timing within a project can limit the application of RRI. Research institutions and funding bodies could provide more practical and sustained support for RRI practice – throughout the research lifespan - including incentives for RRI and clearer RRI/RI roles and career pathways, helping to develop an embedded community of RRI practice throughout the digital research community and beyond.

Keywords: Responsible Research and Innovation; case study; RRI practice; RRI barriers; RRI tools; digital economy; autonomous systems

Introduction

The concept of RRI or Responsible Innovation (RI) is well established as a term in science and technology governance (Von Schomberg, 2011; Stilgoe et al., 2013). The Engineering and Physical Sciences Research Council (EPSRC) has adapted the European vision of RRI, developing the AREA framework (Anticipate, Reflect, Engage and Act) for RRI (Owen et al., 2021).This framework has been expanded within the ICT sector into the AREA Plus framework which elaborates the AREA framework within the "4Ps" (Process, Product, Purpose and People) (Jirotka et al., 2017). Notwithstanding this initiative, the interpretation and adoption of RRI remains fragmented and inconsistently applied. We understand RRI as a dynamic process, an opportunity throughout the lifespan of a project to anticipate and

reflect on the intended and possible unwanted impacts of innovative products, to prevent and mitigate harms and foster socially desirable and trusted outcomes. For that process to be meaningful and sustained, it must be achieved in a collective manner, embracing co-design through stakeholders' and end-users' engagement, which is a pivotal element in RRI.

The relevance of RRI to trustworthy autonomous systems is evident: many systems are already on or close to market, where they apply novel technologies in ways that directly and indirectly impact a multitude of stakeholders including the general public. Centred on a UK University, the research programmes where this work took place have strong commitments to RRI which reflect their respective programmes of work, being particularly focussed on the trustworthiness of autonomous systems and the ethical use of personal data in hybrid (i.e, physical/digital), social, and trusted products. Both research programmes are concerned with emerging digital technologies that are both shaping and being shaped by society. As a result, much of the work within these programmes is organised as shorter focussed research projects that can respond in a timely way to this rapidly changing sociotechnical landscape. This reflective article reports on the experience of one project dedicated to RRI, which started in April 2021 with an initial 12-month duration.

The aim of our project is to explore current RRI practices among researchers linked to these research programmes, to identify barriers and facilitators for RRI implementation and to specifically support researchers in realising RRI within their individual projects. Our project includes: scoping interviews probing stakeholders' understandings of responsible research and current practices of RRI; in depth case studies of RRI with selected research projects; preliminary development of a programme approach to RRI; developing a system to support the creation, use and analysis of card-based tools and activities to support RRI (the [XXX] platform (ref. removed for review)); and support for RRI policy impact. The interviews and case studies aim to identify key RRI factors and strengths/gaps in existing RRI practice, to feed into the co-creation of guidelines for RRI implementation aligned to the research programmes' motivations and challenges. This also supports the co-creation of RRI tools and activities to help researchers and relevant industry partners to identify RRI opportunities and challenges, gain clarity around RRI and provide awareness and guidance on how to translate the EPSRC RRI AREA framework and other RRI-related approaches into embedded RRI practice. Ultimately, this research aims to establish an approach to RRI practice that can provide guidance for current and future research in trustworthy hybrid and digital systems and services and can be generalised to RRI practice within the broader digital sector. The project thus aligns with discipline-based discussions of value-sensitive design (Simon, 2017) and ethics by design (WEF, 2020).

Within the digital technology and ICT environment, historically, emphasis has been given to the anticipation element of RRI to avoid innovation causing harmful consequences to society, stressing that innovation should be ethically designed (Stahl et al., 2013; Dignum et al., 2018). However, actual examples of how this can best be achieved in practice are scarce. More recently, the areas of co-creation (Robinson et al., 2021) and meaningful collaboration (Jarmai & Vogel-Pöschl, 2020) have attracted increasing attention within the RRI ecosystem. However, the barriers and facilitators of co-creation with stakeholders, in particular the

challenges of stakeholder engagement with RRI practice throughout the lifetime of a project, have been discussed in rather abstract terms. In this reflective article we would like to share our experiences and views from our project, with a focus on our engagement with a wide range of stakeholders including researchers, academics, project partners and advisory group members. In particular, this paper focuses on some tensions expressed by participants and lessons learnt from this project so far including: the time and timing required to put RRI into practice, some confusion of what RRI implementation entitles, and the need for institutions to provide more practical support and incentives to aid early and embedded RRI practice within a project.

Methodology

This paper was written during the COVID-19 pandemic, therefore all the research activities were conducted online (in Microsoft Teams). The study was approved by the Research Ethics Committee in the lead author's department. We completed 14 one-to-one semi-structured interviews with researchers, academics, and support staff with different levels of experience, including: three academics (two principal investigators and one co-investigator); eight Research Fellows with different levels of experience; a researcher at an external partner organisation; a final year PhD student; and a senior research facilitator. Participants were recruited from UK research institutions: six of the participants were working on projects specifically related to trustworthy autonomous systems (AS) and eight on other projects linked to the research programmes, four of whom had links to projects on the use of AS. Participants' research backgrounds included: computer science, social science, mental health, psychology, engineering, law, and linguistics. All bar two were working on short-term projects (12 months). During the research interviews, participants were asked: their understanding of responsible research and how they have put it into practice; examples of facilitators and barriers of its implementation within their projects; their experience/ training on RRI; whether and how their institution could better aid responsible research practice. Participants were recruited without regard to their level of knowledge of RRI and were asked to respond the questions according to their thoughts and experiences about conducting responsible research within their project at the time the interview took place.

We engaged with two research projects as example case studies of RRI practice, one from each research programme. Both were short-term projects (12 months). The aim was to co-create an RRI activity plan with each project research team, through a series of activities: 1) research interviews (included in those described above, with 3 or 4 participants from each group); 2) completing the ORBIT RRI online self-assessments (Stahl, 2017; de Heaver et al., 2020); and 3) a reflective team workshop conducted online using a card-based tool (ref. removed for review), to discuss RRI in relation to a technology they were developing in their project. The [XXX] cards are an ethics-by-design tool which seek to highlight questions and encourage reflection and engagement around ethical issues (ref. Removed for review). The card-based workshops took place close to the start of one of the case study projects, and half-way through the other. The workshops focused on enabling the researchers to identify potential ethical benefits and harms of their technology before identifying ways or maximising the benefits and minimising the harms and the practical challenges of the considered approaches.

We have also established an advisory group with nine experts in the field and have so far held two meetings online. The first focussed on facilitators and barriers of RRI and possible gaps in current practice and the second one focussed on assessing RRI for trustworthy hybrid and autonomous systems.

Findings

At the time of writing, our project is ongoing so the reflections presented in this paper are based on our preliminary results and experience from the first part of the project. We have mapped: participants' understandings of R(R)I, including attitudes, tensions and paradoxes; how participants have applied RRI in their current projects; perceived strengths of and barriers to RRI; and opportunities for improvements. Results from the interviews indicate that whilst all participants feel that conducting responsible research is important to them and they appreciate the value of promoting more anticipatory and reflective practice in RI, many currently approach it through the lens of research integrity and research ethics. Most participants had no formal training in RRI (two did during their PhD) but had gained some experience of conducting responsible research in their work. One researcher with background in law, AI and data ethics had a particular interest in RRI. Many participants mentioned that applying RRI within a project takes a lot of time and can be an obstacle to making progress and addressing the project objectives. Although some believed that, especially in short projects, there is limited time to conduct RRI in practice, dedicating time to conduct RRI-focused sessions even within a short project helped participants to plan, anticipate possible socially undesirable outcomes, reflect, and address some of those issues ahead of time before conducting other research activities. Some participants saw the time dedicated to key aspects of RRI within their projects (e.g., anticipation and reflection) as something intrinsic to their research. Other participants mentioned that for researchers to value the investment of time in applying RRI within their project they would require more incentives and support from their departments and institutions. Some believed their institutions should champion the benefits of embedding responsible practice into research in a practical way. Suggestions on how to achieve this included providing research teams with in-person RRI support (e.g., a researcher with expertise in RRI practice) throughout the span of the project to ensure best practice. Participants also suggested their institutions should offer more incentives to researchers to embed RRI practices, e.g., awards for best RRI practice. A participant pointed out that this not only motivates individuals and research teams but would also raise awareness within their research community of the benefits of RRI.

The ORBIT RRI online self-assessment tools are quite broad and comprehensive, covering high-level areas of RRI and ensuring the inclusion of a wide range of issues (including gender and EDI). Participants found these tools relatively time-consuming to complete; they are not specific to digital-technology based innovation, and some elements were felt to be difficult to apply to these projects. Participants also found it difficult to complete some elements that referred to the wider organisation, rather than to their department or research group, or their specific research project.

The card-based workshops were felt to be valuable in enabling, promoting, and providing space for RRI discussions within the project teams. This enabled researchers to anticipate and reflect on ethical and pragmatic questions of significance to their research projects that

may have otherwise gone unconsidered. However, it was less clear how these discussions could lead to action and be effectively integrated into the wider work of the project, warranting further consideration and development. This part of the work has been significantly influenced by the necessity of conducting all the workshops online (due to COVID-19), which resulted in a very different workshop experience compared to face-to-face card workshops (ref. removed for review).

Discussion with the advisory group has provided insightful discussions and helped to reflect on and shape our research activities. The remaining work of the project is to integrate the findings from all these research activities into a co-designed approach to RRI for these research programmes.

Discussion

The [name] project has begun to elucidate the complexity and importance of RRI in this particular context, which provides valuable groundwork for future work around RRI/RI, both in and beyond these research programmes. In this section we highlight our observations about engaging with responsible research practice in the individual research projects that constitute larger research programmes with explicit concerns for trustworthy and ethical hybrid and digital systems; we consider in turn: when RRI is "done", how RRI is understood, and how it can be supported.

In terms of when RRI activities take place, we experienced logistical challenges engaging with a project for a case study that was already quite advanced. RRI arguably gains much of its value from anticipating possible undesirable consequences of research and allowing space for reflection, engagement, and action. When projects have already been designed, funded and are well underway it may be too late for RRI interventions to significantly influence them. Even otherwise valuable RRI activities, for example the [XXX] cards, may have little scope to lead to practical action within the lifetime of the project, which can cause frustration with the time and resources allocated to these activities. The reflections from these activities may be more beneficial to planning of later projects, especially if the same researchers will be involved.

The sense of time pressure is particularly acute for short research projects (e.g., average of 12 months long), as shared by most participants, where RRI can be seen as something which takes significant time from other time-critical aspects of a project. It is very difficult to find a good way to balance the project's various research commitments (e.g., technical development, studies, data collection and analysis, deliverables), while also embedding RRI practices. Building time for RRI into the proposal development process can feel even more challenging, as proposal development often has no explicit time or resources allocated to it.

When it comes to applying RRI it is often unclear to researchers what the expectations of funding bodies and academic institutions are. Within a research organisation or project RRI is often represented and understood in an instrumental manner as something to be "done", rather than as an attitude or approach to conducting research. Some participants described carrying out work that is in line with RRI principles without necessarily making an explicit

link to RRI. This perhaps originates from equating RRI with functional elements of project administration and governance which can be seen, demonstrated, and monitored, such as the identification of RRI with the research ethics application process or the need to provide an explicit RRI statement in funding proposals (as was the case in these research programmes). While these are important elements of RRI practice they are not likely to address the breadth of potential social and ethical considerations. UKRI funded institutions have recently introduced mandatory RRI sections on funding applications; we believe that this is a valuable prompt for research teams to anticipate and reflect on relevant issues when designing a project. However, there is currently little or no formal institutional support or follow up in this area. This can easily lead to a "tick box" mindset which perceives RRI only as an organisational barrier to be overcome or bypassed.

We also found conflicting views and tensions on whether and how RRI practice should be evaluated or measured, which was also highlighted by the advisory group. Institutional management practices often revolve around quantification and measurement, for example the assessment of organisational strategy and improvement in terms of Key Performance Indicators. In the context of RRI, "measuring" anticipation or reflection seems contrary to the principles these elements are intended to embody, in that quantification always involves a degree of reductionism, rendering specific aspects visible or invisible. Whether RRI implementation can and should be "measured", and if so how to prove success and long-term impact, are real challenges and open questions within the RRI ecosystem (Stahl et al., 2021). There is an inherent tension in that the elements of the AREA framework are less amenable to "measurement", yet without mechanisms of accountability RRI may be less likely to be recognised, valued, and understood amongst the competing demands for a researcher's time. A broader institutional appreciation of RRI, addressing researchers' needs and their desire for ways of recognising RRI within their projects, may help to mitigate this tension. However, it may be harder to demonstrate and provide evidence of having applied responsible research within a project to the benefit of society.

There is therefore a particular need for RRI to clearly demonstrate its value to researchers in terms of "what's in it for me?" beyond simply getting the research done. One strategy here may be to look for and develop RRI-specific research contributions of each individual project (or group of projects). This is to say, RRI processes can be valued research activities in their own right within a project, with outcomes (e.g., publications, co-created products and/or processes, policy impact) that clearly reward the additional effort invested by the research team. This may also form the basis of fruitful collaboration and knowledge sharing between RRI specialists and other researchers/innovators.

Turning finally to possible support for RRI practice, we consider first the role of tools. Much of the work of this project has made us of tools (the ORBIT RRI tools and [xxx] cards). However, simply asking researchers to use these tools is not enough to make responsible research and innovation "happen". These tools are typically intended to be flexible and widely applicable in use, so careful selection and significant support may be important in making them valuable across contexts. In our case the ORBIT Organisation Self-Assessment would have been better completed once at the organisational or research programme level

rather than for each project. There can also be tensions between the perceived and intended functions of these tools. For example, using the [XXX] cards may be perceived by some researchers as a way of being able to say they have "done" RRI, whereas this tool is designed to raise questions and promote conversations, and the "doing" of RRI is at least as much in this conversation, team discussion (e.g., anticipation and reflection) and the actions taken as a result. In our experience, a skilled facilitator can help greatly in this process.

Within the context of a larger research programme or centre that facilitates smaller projects, the programme or centre itself can address many of the challenges of timing and resourcing RRI. One element might be to consider – and support – the principles of RRI in advance of the conception, design and submission of any research proposal and idea. Resources dedicated to RRI (such as our own project) could be provided in advance to support researchers in shaping their proposal to include meaningful consideration of RRI challenges before and throughout each project. Ideally, this prior anticipation, reflection, and engagement around a potential project would result in a project designed to meet the goals of responsible innovation. Whilst such support would improve the proposal and its chances of success, provision of RRI support for proposals that are then unsuccessful may feel like a waste of time and resources. However, we argue that such effort would not be wasted, as the insights and reflections would often help to shape future projects and proposals. Once projects are funded, ongoing support, anticipation, and reflection should be a sustained and dynamic process throughout each project, helping to align it to socially desirable outcomes. This might include dedicated support for RRI roles throughout the lifetime of research projects.

Finally, although larger funded projects and research centres have some ability to underwrite RRI activities, we note that broader support for RRI requires a significant commitment at institutional levels, including funding bodies such as UKRI. Our views from this project align with RRI-Practice Policy Recommendations (Owen et al., 2019) aimed at strengthening RRI practice (in our case from a bottom-up rather than a top-down perspective to RRI implementation). We believe that a pervasive commitment to and support of RRI at an institutional level would be needed to establish a strong research culture of RRI, sustainable employment opportunities, and career development pathways for RRI/RI specialists.

Conclusions

In this short article we have presented some of our experiences and reflections to date from a project focussed on RRI practice within two linked programmes of research addressing trustworthy autonomous systems and the ethical use of personal data. Even with immediacy of possible social impacts and clear direction from the leaders of these research programmes in terms of promoting RRI, researchers often experience uncertainty and a sense of time pressure when embedding RRI into specific projects, especially projects with short timescales. Engagement with key RRI actors including researchers, academics, research support staff and external partners is essential to identify and address possible challenges within RRI practice. Many researchers would welcome more practical support for

RRI from the early stages of proposal development, as well as at later points in the lifetime of the project. While tools have a useful part to play, in our experience their value is strongly enhanced by the involvement of research staff with expertise in RRI. Developing and sustaining such expertise may ultimately depend on clearer institutional valuing of RRI in practice, for example explicitly rewarding RRI and establishing career paths for RRI specialists.

Disclosure statement
No potential conflict of interest was reported by the authors.

Acknowledgements
Section removed for anonymous review.

References

de Heaver, M., Jirotka, M., Nulli, M., Stahl, B. C., & Holter, C. T. (2020). RRI intensity: A proposed method of assessing the requirement for responsible innovation in ICT projects. In *Assessment of Responsible Innovation* (pp. 297–315). Routledge.

Dignum, V., Baldoni, M., Baroglio, C., Caon, M., Chatila, R., Dennis, L., Génova, G., Haim, G., Kließ, M. S., Lopez-Sanchez, M., Micalizio, R., Pavón, J., Slavkovik, M., Smakman, M., van Steenbergen, M., Tedeschi, S., van der Toree, L., Villata, S., & de Wildt, T. (2018). Ethics by Design: Necessity or Curse? *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 60–66. https://doi.org/10.1145/3278721.3278745

Jarmai, K., & Vogel-Pöschl, H. (2020). Meaningful collaboration for responsible innovation. *Journal of Responsible Innovation*, *7*(1), 138–143. https://doi.org/10.1080/23299460.2019.1633227

Jirotka, M., Grimpe, B., Stahl, B., Eden, G., & Hartswood, M. (2017). Responsible research and innovation in the digital age. *Communications of the ACM*, *60*(5), 62–68. https://doi.org/10.1145/3064940

Owen, R., Forsberg, E.-M., & Shelley-Egan, C. (2019). *RRI-Practice Policy Recommendations and Roadmaps, RRI-Practice project report. Deliverable 16.2*.

Owen, R., Pansera, M., Macnaghten, P., & Randles, S. (2021). Organisational institutionalisation of responsible innovation. *Research Policy*, *50*(1), 104132. https://doi.org/10.1016/j.respol.2020.104132

Robinson, D. K. R., Simone, A., & Mazzonetto, M. (2021). RRI legacies: Co-creation for responsible, equitable and fair innovation in Horizon Europe. *Journal of Responsible Innovation*, *8*(2), 209–216. https://doi.org/10.1080/23299460.2020.1842633

Simon, J. (2017). Value-Sensitive Design and Responsible Research and Innovation. In S. O. Hansson (Ed.), *The Ethics of Technology: Methods and Approaches* (pp. 219–236). Rowman & Littlefield International.

Stahl, B. C. (2017). The ORBIT Self-Assessment Tool. *The ORBIT Journal*, *1*(2), 1–12. https://doi.org/10.29297/orbit.v1i2.59

Stahl, B. C., Akintoye, S., Bitsch, L., Bringedal, B., Eke, D., Farisco, M., Grasenick, K., Guerrero, M., Knight, W., Leach, T., Nyholm, S., Ogoh, G., Rosemann, A., Salles, A., Trattnig, J., & Ulnicane, I. (2021). From Responsible Research and Innovation to responsibility by design. *Journal of Responsible Innovation*, *8*(2), 175–198. https://doi.org/10.1080/23299460.2021.1955613

Stahl, B. C., Eden, G., & Jirotka, M. (2013). Responsible Research and Innovation in Information and Communication Technology: Identifying and Engaging with the Ethical Implications of ICTs. In *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (edited by R. Owen, J. Bessant and M. Heintz, pp. 199–218). John Wiley & Sons, Ltd. 10.1002/9781118551424.ch11

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, *42*(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Von Schomberg, R. (2011). *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields* (SSRN Scholarly Paper ID 2436399). Social Science Research Network. https://papers.ssrn.com/abstract=2436399

WEF. (2020). *Ethics by Design: An organizational approach to responsible use of technology (White Paper)*. World Economic Forum. https://www.weforum.org/whitepapers/ethics-by-design-an-organizational-approach-to-responsible-use-of-technology/

Formatted: Normal

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Formatted: Font: (Default) Times New Roman

Formatted: Normal