This paper was presented at The ISPIM Innovation Conference – Innovating Our Common Future, Berlin, Germany on 20–23 June 2021.
Event Proceedings: LUT Scientific and Expertise Publications: ISBN 978-952-335-467-8

Responsible innovation in Multidisciplinary Research and Innovation Projects: Moving from Principle to Practice

Stergios Aidinlis*

Trilateral Research Ireland, 2nd Floor, Marine Point, Belview Port, Waterford, X91 W0XW, Ireland
E-Mail: stergios.aidinlis@trilateralresearch.com

Agata Gurzawska

Trilateral Research Ireland, 2nd Floor, Marine Point, Belview Port, Waterford, X91 W0XW, Ireland
E-mail: agata.gurzawska@trilateralresearch.com

* Corresponding author - disclaimer: both authors contributed equally to the writing of the paper.

Abstract:

The process of translating and operationalising responsible innovation principles is crucial for ensuring that they can directly impact the architectural choices made by technology developers in the innovation process, particularly at the stage of design. This paper focuses on a particular type of innovation activity: multidisciplinary research and innovation (R&I) projects involving data-driven technology development and requiring the participation of multiple disciplinary research backgrounds, often members of international publicly funded consortia. Through our proposed approach, which we frame as the *Trilateral TouchPoint Risk Assessment Table™*, we seek to provide an applied and practical toolkit for Social Sciences and Humanities (SSH) researchers to enable them to navigate their relationship with engineers and effectively embed responsible innovation, ethical legal and societal aspects (ELSA) by-design. In doing so, we draw on our experience of applying the tool in two EU-funded Horizon2020 security research projects: COPKIT and DARLENE.

Keywords: Responsible innovation, multidisciplinary R&I, innovation management, technology assessment (TA), impact assessment (IA), ELSA.

1 Introduction

While there is a rich literature and academic discourse about embedding responsible innovation in technology development, the very process of

Event Proceedings: LUT Scientific and Expertise Publications: ISBN 978-952-335-467-8

translating/operationalising ethical, legal and societal values so that they can directly impact the architectural choices made by technology developers at the stage of design is relatively underexplored. In this paper, we focus on a particular type of innovation activity: multidisciplinary research projects involving data-driven technology development and requiring the participation of multiple disciplinary research backgrounds, often members of international publicly funded consortia. Considering the time pressures and the large number of actors involved in multi-disciplinary technology development research, ethicists, lawyers and social scientists, hereafter referred to as Social Sciences and Humanities (SSH) researchers, may struggle on how to navigate their relationship with engineers to effectively embed responsible innovation and ethical, legal and societal aspects (ELSA) by-design.

Deriving from existing methodologies and approaches to responsible innovation such as technology assessment, responsible research and innovation (RRI), impact assessment (IA) and design-thinking, this paper investigates how SSH researchers can communicate effectively with engineers within multidisciplinary technology development research projects to directly shape the architecture of technology and embed fundamental values such as ethical principles and privacy at the stage of design. As a result, we present our emerging methods framework for pursuing responsible innovation by-design in multidisciplinary research projects, which we frame as the *Trilateral TouchPoint Risk Assessment Table™*.

This paper is structured as follows: Sections 2 and 3 provide an overview of the specificities and challenges engendered in multidisciplinary research projects, as well as existing methods and approaches to practicing responsible innovation, ethical, legal and societal aspects (ELSA) in the face of such challenges. Section 4 presents the *Trilateral TouchPoint Risk Assessment Table™* and discusses how the tool was used in two H2020-funded projects in the context of data-driven law enforcement, namely COPKIT and DARLENE. Section 5 summarises our findings and reflects on areas for improvement.

2. Specificities and challenges of multidisciplinary R&I projects

The literature on innovation management shows that integrating distant knowledge often yields novel and more innovative outcomes than working in a close domain (Hacklin & Vallin, 2013). Within the broader innovation landscape, multidisciplinary technology development research and innovation (R&I) projects, often carried out by large international consortia, provide us with a particularly interesting case for responsible innovation and ELSA-by-design. On the one hand, these projects are, in most cases, publicly funded and require the collaboration between technology developers and SSH researchers with a view to promoting research into the ethical, legal or/and societal impacts of the developed technologies (European Commission, 2020). Academics and policy makers often refer to the concept of Responsible Research and Innovation (RRI), which embodies the innovation process and its marketable products that are ethically

acceptable, socially desirable and respond to the needs and expectations of people and the society (Von Schomberg, 2013). Four main characteristics of RRI include: (1) inclusion (also called engagement or involvement of society), (2) anticipation (assessment at an early stage in research and innovation (R&I) of benefits and risks), (3) reflexivity (reflecting on values and beliefs during R&I), and (4) responsiveness (the ability to change routines, structures and systems to adapt to changing circumstances and new insights) (Owen et al., 2013). One of the most well-known definitions of RRI emphasises the *interactive* character of the R&I process, where societal actors and innovators become mutually responsive to each other (Von Schomberg, 2013). Various other definitions of RRI exist and numerous authors have sought to operationalise it in different ways (Sutcliffe, 2011; Rip, 2014; Taebi et al. 2014; Foley et al., 2016; Lindner et al., 2016; Burget et al., 2017; Ribeiro et al., 2017; Reber, 2018). In that sense, one might think that such R&I projects provide very fertile ground for approaches that can effectively shape architectural choices at the stage of technology design.

On the other hand, however, the practical experience of participating in such projects indicates the complexity of seeking to embed e.g., responsibility and ELSA-by-design. There is a plethora of participating and affected stakeholders that need to collaborate, or at least be consulted, in these projects. Technology developers, SSH researchers, public authorities and industry bodies who are often the addressees/real-world users of the developed technologies, civil society and academic organisations who are interested in the subject matter of research. In some cases, there might be overlaps between the previous categories: law enforcement agencies (LEAs) are public authorities, but they might also participate in a technology development capacity within a project (COPKIT Project, 2020).

From the perspective of SSH researchers, seeking to embed responsibility and ELSA at the stage of technology design requires not only effective communicative strategies that bridge disciplinary divides with technology developers, but also adequate consideration of the, potentially conflicting, interests and concerns of these numerous stakeholders. There are inherent epistemological difficulties in assessing the relative validity of normative claims about the implications of future technologies. As a result, SSH researchers and engineers may not be on the same page regarding the ethical, legal and societal implications of an emerging technology and, thus, the needs to ember certain considerations into the design stage (Groves, 2015).

This becomes even more challenging when considering the time pressures of multidisciplinary research projects, where funders normally expect very ambitious dissemination plans of the technological findings in a way that will tangibly generate policy impacts within very strict deadlines (European Commission, 2021). Day-to-day realities of developing emerging technologies involve significant time pressures and require very efficient and prompt communication between SSH researchers and engineers, also made difficult by the lack of sufficient integration of responsible innovation and ethics in the day-to-day work of engineers (Steen, 2015).

Event Proceedings: LUT Scientific and Expertise Publications: ISBN 978-952-335-467-8

Furthermore, a crucial characteristic of innovation is that it is not a single action, idea, or invention of a single new device but rather an integrated process involving various phases (Conway & Steward, 2009), though not necessarily linear including various feedback-loops (Eveleens, 2010). The innovation process involves a high level of uncertainty and risk, thus some authors refer to the *innovation journey* (Van der Ven & Poole, 1990). Integration of responsibility and ELSA into R&I multidisciplinary projects needs models and prescriptions that help SSH researchers to deal with the challenges posed by non-linearity, uncertainty and multidisciplinarity in an agile manner with feedback, slack, and direction to develop fruitfully (Hacklin & Wallin, 2013). The complexity of emerging technologies requires the adoption of more flexible and versatile responsible innovation and ELSA-by-design methodologies, also considering the particularities of different technological architectures across research projects. There is, thus, no one-size-fits-all solution, unless that solution can adjust to different needs and requirements based on the project. Hence, the main challenge in multidisciplinary R&I projects relates to embedding the ELSA that need to be considered from an SSH research perspective within the innovation process and throughout its lifecycle, not only providing one-off assessments at certain stages of a project. Approaches at the intersections of technology assessment, impact assessment, RRI and design-thinking, seek to address these challenges in multidisciplinary collaborations.

3. Responsible innovation & ELSA by design – outlining the field

With the aim of ensuring responsibility and ELSA-by-design, R&I projects may draw on a variety of approaches and methodologies. The field of innovation management brings technology assessment (TA), which aims to evaluate potential, and actual, impacts of new technologies on industry, the environment and society; and to develop instruments to steer technology development in more desirable directions (Swierstra & Rip, 2007; Tran & Daim, 2008; Brey 2012). The assessment is based on known or potential applications of the technology, taking into consideration consequences that are unintended, indirect or delayed (Nielsen et al., 2017). Moreover, multidimensional R&I projects may involve a range of assessment methods such as impact assessment (IA), i.e., a process of identifying the future consequences of a current or proposed action (IAIA). Some examples include effects on environment (environmental IA), society (social or societal IA), health (HIA) or human rights (HRIA). Furthermore, RRI emphasises that innovation should be analysed and assessed with the goal of influencing innovation processes to make them more ethical (Gurzawska, 2021). Such approaches as ethical technology assessment (eTA) and ethical impact assessment (eIA) respond to this need. Reijens et al. (2017) identify three types of methods for practising ethics in R&I: *(1) ex ante methods, dealing with emerging technologies, (2) intra methods, dealing with technology design, and (3) ex post methods, dealing with ethical analysis of existing technologies.* This division touches upon the question

of various phases of the innovation process and the fundamental need for agility in RRI.

'Design-thinking' approaches also exemplify the agile methodological underpinnings of RRI. These approaches perceive design as a cognitive process, a 'more interpretative, intuitive mind-set' (Bason, 2010), that extends beyond the design of products and can be applied more broadly as a way of thinking. While not synonymous with 'user-centred design' (Dorst, 2011), design-thinking prioritise the perspectives of system 'users', seeking to iterate continuously between their needs and the technology or product under development (Kimbell, 2012), assessing the alignment of the latter with the former. In the case of multidisciplinary technology development projects (Goldsworthy and Ellam, 2019), the 'users' must not be strictly confined to end-users of the technology, but also, and crucially, with the society at large and particular social groups (e.g., minorities, disabled individuals etc). In that sense, design-thinking suggests that researchers try to empathise with their needs in defining the research problems and testing adequate solutions, with a view to produce a technology that has considered relevant needs and the fundamental values that need to be protected from the outset.

Various other forms of integrating responsibility into the innovation process exist, such as social design, socially responsible design (SRD), design for values, open innovation and eco-design. In this paper, we define responsible innovation as the innovation process and its outcomes that accounts for various forms of responsibility including ethical, legal (including human rights and privacy), societal and environmental aspects and impacts of multidisciplinary R&I projects. Considering the plethora of existing approaches, we now turn to our proposed tool as a practical solution to the challenges of responsible innovation and ELSA-by-design.

4. Proposed tool for practicing RI & ELSA-by-design: *Trilateral TouchPoint Risk Assessment Table™* and Case Studies

In multidisciplinary EU-funded R&I projects, we use the *Trilateral TouchPoint Risk Assessment Table™* to identify and analyse potential ELSA and risks in the entire life cycle of the project and various phases of the innovation process, thus throughout the project's Work Packages (WPs) and tasks. Our approach provides a concrete and applied understanding and management of responsible innovation by operationalising RRI principles and integrating various forms of TA, IA and design-thinking. The potential of our approach lies in its: (1) flexibility, due to its capacity to be adjusted to meet the needs of different technological designs, (2) diversity and inclusiveness since it can bring together effectively different types of expertise within the project, as well as advisory board and potentially feedback from the public and other relevant stakeholders, (3) anticipation, reflection and agility, since it involves sustained dialectic engagement and numerous iterations between the SSH researchers and the engineers throughout the project, and (4) systematicity due to its capacity to

Event Proceedings: LUT Scientific and Expertise Publications: ISBN 978-952-335-467-8

integrate different risk assessment methodologies, covering ELSA of technology development.

Essentially, the *Touchpoint table* consists of four columns. The first column identifies the task to be developed with a brief description of what the task is about. The second column identifies the potential ELSA that might arise during or from the task, their likelihood and impact. The third column identifies an agreement with the task leader about how to manage the issues. The fourth column contains the views of an external ethics advisory board (EAB) regarding the adequacy of the solutions proposed in the third column.

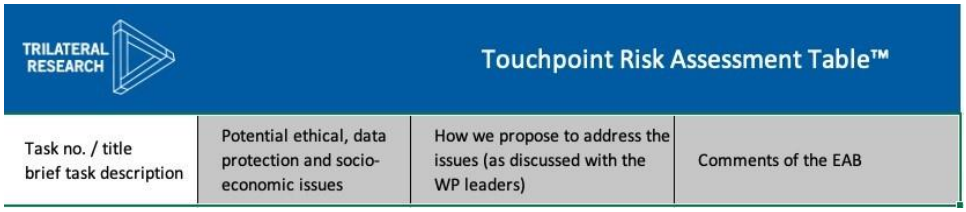

Figure 1 An indicative illustration of the *Trilateral TouchPoint Risk Assessment Table™*

This basic table (Table 1) can be and has been tweaked according to the exigencies of the project being evaluated, as described in the following section, but the above columns are the core of any *Touchpoint table.* The application of our tool is illustrated in the following two case studies.

COPKIT

The COPKIT project aims to support LEAs in analysing, preventing, investigating and mitigating the use of new information and communication technologies by organised crime (COPKIT project, 2020). COPKIT aims to develop Natural Language Processing-based tools focused on analysis of darknet advertisements markets as online digital platforms for illicit activity, such as weapons, drugs, and Crime-as-a-Service (CaaS) related entities such as carding, hacking, etc.

In the COPKIT project, the *Touchpoint table* was used as part of an integrated ethical, legal (including human rights and privacy) and societal impact assessment of the COPKIT ecosystem, technologies and their applications. *The Touchpoint table* helped to assess the benefits of the COPKIT technologies as well as their potential negative impacts (risks) and mitigation strategies. Following the principles of responsible innovation (anticipation and reflection), design-thinking and agility of the innovation process, the impact assessment was divided into three steps with multiple feedback-loops integrated within the innovation process: (1) impact assessment (IA) by SSH partners involving ethicists and lawyers; (2) regular mutual learning and joint IA via dialogue sessions with technical partners (SSH and technical partners); (3) consultation with the ethical and societal impacts advisory board (ESIAB).

For COPKIT purposes, the *Touchpoint table* focused on the impact assessment of COPKIT tools. They were analysed and assessed taking into consideration both innovation process during the lifespan of the project (internal impact) and its outcomes, when potentially used by LEAs after the project ends (external impact).

Furthermore, both positive and negative impacts (risks) were identified to indicate the extent to which the COPKIT tools enhances or undermines European values and principles. The *Touchpoint table* involved an integrated IA including two parts: (1) ethical, legal and societal impact assessment (ELSIA); (2) data protection impact assessment (DPIA). Potential impact and the assessment criteria were derived from the COPKIT framework comprising ELS requirements relevant for the COPKIT technologies, such as fundamental rights and freedoms, data protection regulations and social values identified for the EU and various European countries. Moreover, ethical principles and guidelines were considered including research ethics principles, police ethics standards and those of related to ethics of technology, e.g., the High-Level Expert Group on Artificial Intelligence (HLEG AI). An interesting tweak of *the Touchpoint table* was an application of the CEN Workshop Agreement on ethical impact assessment (EIA) developed as part of the SATORI project (2018). This extension resulted in additional analysis and evaluation categories, namely: (1) the likelihood of impacts to occur; (2) the relative intensity of impacts; (3) potential or actual rights, principles and value conflict and means of resolving it; (4) mitigation strategies; (5) actor(s) responsible for implementing these strategies; and (6) re-evaluation of likelihood and intensity of risk after mitigation strategies.

DARLENE

DARLENE, standing for Deep Augmented Reality Law Enforcement Ecosystem, aims to combine and advance Augmented Reality (AR) and Artificial Intelligence (AI) technologies to enhance the situational awareness of law enforcement officers in the field, particularly in time-critical and life-threatening scenarios (DARLENE Project, 2020). Leading the ethics, legal and societal efforts in the DARLENE project, the *Touchpoint table* has enabled us to introduce a multi-layered and systematic oversight framework that follows the lifecycle of the DARLENE technology (Aidinlis, 2021). Our approach involves sustained dialogue and exchange of information between technology developers and SSH experts, in the interest of achieving accountability in creating innovative AR tools from a responsible innovation perspective. Regular meetings are held between not only SSH experts working for different partners, but also between these experts and technology developers, as well as the project's Ethics Advisory Board, to foster an iterative, collaborative way of thinking about ongoing challenges. Like the common approach in design-thinking (Rowe, 1987), these communicative channels also enable the brainstorming around emerging problems and the prototyping of solutions to these problems. In some cases, technology developers and SSH experts might find that common solutions can address different problems faced by consortium partners.

To provide some examples of embedding responsible innovation principles in the development of DARLENE technology, one common challenge for tech and SSH experts related to the existing machine learning (ML) training datasets for violent acts in the context of computer vision that will enable augmented reality. From a technology perspective, the challenge of *occlusion* emerged, I.e., in some

Event Proceedings: LUT Scientific and Expertise Publications: ISBN 978-952-335-467-8

cases the figures of certain individuals were partially or wholly blocked by other individuals or intervening objects. From an SSH perspective, the challenge of *bias* emerged, i.e., in the existing datasets it would often be the case that the included individuals would tend to predominantly belong to a specific race. After thorough communication between tech and SSH experts, a common strategy emerged: *data augmentation,* i.e., the artificial re-balancing of the existing datasets through a machine-learning algorithm that substitutes the place of figures within the images of the dataset to avoid both occlusion problems and mitigate the risks of bias (e.g., by seeking to produce a final dataset where race percentages are more balanced). Beyond data augmentation, technology developers and SSH experts have fruitfully collaborated within our methodology to address other well-known ethical, legal and societal issues in multidisciplinary data-driven research. A common challenge relates to the difficulty to reconcile the data-intensive nature of intelligent law enforcement tools with data protection and minimisation principles, with a view to scrutinising executive power in this very sensitive context. In that regard, the DARLENE architecture envisages *real-time* and *in-situ* data processing by the smart glasses, avoiding storing or further processing of any data unless this is strictly necessary for the detection, prevention or prosecution of terrorist or criminal activity.

The Touchpoint table was a crucial analysis, documentation, guidance and monitoring mechanism for the agile approach, where ELSA were integrated hand in hand with the dynamic technical development of COPKIT and DARLENE tools. Potential end-users (LEAs) can be reassured that the project partners have considered and addressed ELSA throughout the entire project life cycle (i.e., from requirements gathering and design to integration, demonstration and testing). We believe that such approach contributes towards the success of uptake, supports the identification of stronger sustainability opportunities and ensures increased stakeholder and public trust in the technologies and practices produced by projects like COPKIT and DARLENE.

5. Conclusions and reflections

Our findings bear tangible practical implications for both innovation management professionals and researchers working on the intersections of data-driven technologies and ethics, as well as on research funders in the area of responsible innovation. In the case of professionals and researchers, our findings provide insights into a concrete and applied methodology for translating ethical considerations into technology design requirements. Through drawing on our methodology and the cases where it has been applied, such professionals may benefit and draw connections with their own work. In the case of funders, our findings may stimulate thinking about improving current approaches to enhance collaboration between SSH researchers and engineers in large multidisciplinary projects.

Engagement with these audiences is important for the purposes of reflecting for the potential improvement of our tool regarding the most efficient communicative

channels between engineers and SSH researchers, the improvement of methods to include numerous stakeholders in the impact and risk assessment process and balancing their often-competing interests, as well as the limitations in applying the tool in other RRI contexts.

References and Notes

Aidinlis, Stergios (2021), 'Designing a multi-layered ethical and legal oversight for developing augmented reality technology in DARLENE' https://www.darleneproject.eu/designing-a-multi-layered-ethical-and-legal-oversight-for-developing-augmented-reality-technology-in-darlene/

Bason, Christian (2010), *Leading Public Sector Innovation: Co-creating for a Better Society*, Bristol: Policy Press, p 138

Brey, P.A., (2012). Anticipatory ethics for emerging technologies. *Nanoethics*, *6*(1), pp.1-13.

Burget, M., Bardone, E. and Pedaste, M., (2017). Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and engineering ethics*, *23*(1), pp.1-19.

Conway, Steve; Steward, Fred (2009), "Managing and shaping innovation", Oxford University Press, p. 10.

COPKIT Project (2020), 'Organizations and Technologies: Bridging the Gap Between Law Enforcement Needs and Industry Solutions' (2020) https://copkit.eu/organizations-and-technologies-bridging-the-gap-between-law-enforcement-needs-and-industry-solutions/

DARLENE Project (2020), 'About DARLENE Project' https://www.darleneproject.eu/

Dorst, Kees (2011) 'The core of "design thinking" and its application' Design Studies, 32(6)

European Commission (2020), 'Horizon 2020' (2020) https://ec.europa.eu/programmes/horizon2020/

European Commission (2021), 'Implementation strategy for Horizon Europe' https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_implementation-strategy_he.pdf

Eveleens, C., (2010). Innovation management; a literature review of innovation process models and their implications. *Science*, *800*(2010), pp.900-916.

Foley, R.W., Bernstein, M.J. and Wiek, A., (2016). Towards an alignment of activities, aspirations and stakeholders for responsible innovation. *Journal of Responsible Innovation*, *3*(3), pp.209-232.

Goldsworthy, Kate ; Ellams, Dawn (2019) Collaborative Circular Design. Incorporating Hacklin, Fredrik; and Wallin, Martin W., (2013). Convergence and interdisciplinarity in innovation management: a review, critique, and future directions. *the service industries journal*, *33*(7-8), pp.774-788.

IAIA (International Association for Impact Assessment), https://www.iaia.org/

Life Cycle Thinking into an Interdisciplinary Design Process (2019), The Design journal, 01 April 2019, Vol.22, pp.1041-1055

Lindner, R., Kuhlmann, S., Randles, S., Bedsted, B., Gorgoni, G., Griessler, E., Loconto, A. and Mejlgaard, N., (2016). Navigating towards shared responsibility in research and innovation: approach, process and results of the Res-AGorA Project.

Groves, Christopher (2015) 'Logic of Choice or Logic of Care? Uncertainty, Technological Mediation and Responsible Innovation' 9 Nanoethics 321-333
Gurzawska, A., 2021. Responsible Innovation in Business: Perceptions, Evaluation Practices and Lessons Learnt. *Sustainability*, *13*(4), p.1826.
Kimbell, Lucy (2012) 'Rethinking design thinking: Part 1' Design and Culture, 3(3), 285–306
Nielsen, R.Ø.; Gurzawska, A.; Brey, P. (2015). Principles and Approaches in Ethics Assessment Ethical Impact Assessment and Conventional Impact Assessment https://satoriproject.eu/media/1.a-Ethical-impact-assessmt-CIA.pdf
Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E. and Guston, D., (2013). A framework for responsible innovation. *Responsible innovation: managing the responsible emergence of science and innovation in society*, *31*, pp.27-50.
Reber, B., (2018). RRI as the inheritor of deliberative democracy and the precautionary principle. *Journal of Responsible Innovation*, *5*(1), pp.38-64.
Ribeiro, B.E., Smith, R.D. and Millar, K., (2017). A mobilising concept? Unpacking academic representations of responsible research and innovation. *Science and engineering ethics*, *23*(1), pp.81-103.
Rip, A. (2014). "The past and future of RRI." Life Sciences, Society and Policy **10**(1): 17.
Rowe, Peter G. (1987) *Design Thinking*, Cambridge, MA: MIT Press
Sutcliffe, Hilary, (2011). A report on responsible research and innovation. *MATTER and the European Commission*.
SATORI project (2018), https://satoriproject.eu/events/event/ethical-impact-assessment-framework-for-research-and-innovation/
Steen, Marc (2015) 'Upon opening the black box and finding it full: Exploring the ethics in design practices' Science, Technology, & Human Values, 40(3), 389-420.
Swierstra, T. and Rip, A., 2007. Nano-ethics as NEST-ethics: patterns of moral argumentation about new and emerging science and technology. *Nanoethics*, *1*(1), pp.3-20.
Taebi, B., Correlje, A., Cuppen, E., Dignum, M. and Pesch, U., 2014. Responsible innovation as an endorsement of public values: The need for interdisciplinary research. *Journal of Responsible Innovation*, *1*(1), pp.118-124.
Tran, T.A. and Daim, T., 2008. A taxonomic review of methods and tools applied in technology assessment. *Technological Forecasting and Social Change*, *75*(9), pp.1396-1405.
Van de Ven, A.H. and Poole, M.S., (1990). Methods for studying innovation development in the Minnesota Innovation Research Program. *Organization science*, *1*(3), pp.313-335.
Von Schomberg, Rene, (2013). A vision of responsible research and innovation. *Responsible innovation: Managing the responsible emergence of science and innovation in society*, pp.51-74.