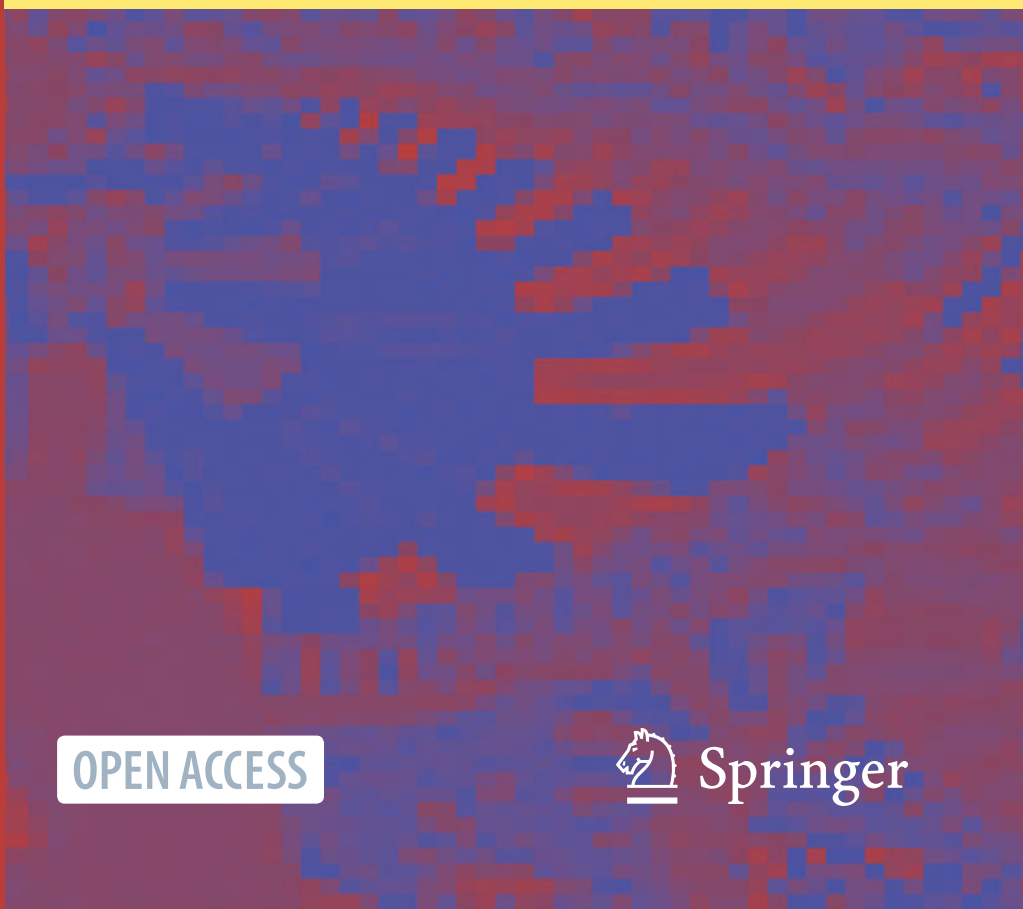
State-of-the-Art Survey

LNCS 13875

Elsa González-Esteban
Ramón A. Feenstra
Luis M. Camarinha-Matos (Eds.)

Ethics and Responsible Research and Innovation in Practice

The ETHNA System Project



OPEN ACCESS

Springer

Lecture Notes in Computer Science **13875**

The series Lecture Notes in Computer Science (LNCS), including its subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), has established itself as a medium for the publication of new developments in computer science and information technology research, teaching, and education.

LNCS enjoys close cooperation with the computer science R & D community, the series counts many renowned academics among its volume editors and paper authors, and collaborates with prestigious societies. Its mission is to serve this international community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings. LNCS commenced publication in 1973.

Elsa González-Esteban · Ramón A. Feenstra ·
Luis M. Camarinha-Matos
Editors

Ethics and Responsible Research and Innovation in Practice

The ETHNA System Project

Springer

Editors
Elsa González-Esteban
Universitat Jaume I
Castelló, Spain

Ramón A. Feenstra
Universitat Jaume I
Castelló, Spain

Luis M. Camarinha-Matos
NOVA School of Science and Technology,
UNINOVA - CTS and LASI NOVA
University of Lisbon
Monte da Caparica, Portugal

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-031-33176-3 ISBN 978-3-031-33177-0 (eBook)
https://doi.org/10.1007/978-3-031-33177-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Innovation and research centers, universities, and their researchers need to achieve results in very competitive contexts, as they are increasingly subject to high-pressure situations, which can lead to some unacceptable behaviors. At the same time, there is a growing awareness of the need to conduct such research and innovation activities with honesty and integrity, respecting well-accepted practices and shared ethical and social values.

In this context, there has been a growing discussion about the responsibility that institutions have in relation to promoting policies conducive to research integrity, public engagement, open access, and gender equality in research processes. Various ethics management instruments – such as codes of ethics and best practices or ethics committees – can play a prominent role in achieving the above goals. In recent years, ethics committees, for example, have been gradually consolidated in different contexts, and it is not uncommon for specific funding organizations or specific scientific journals to make it a prerequisite for funding or publishing research work. However, the implementation and consolidation of these ethics tools have undergone a long and complex process. Moreover, it is also not a homogeneous process in each of the different countries. Even when institutions attempt a top-down implementation approach, individual researchers and other stakeholders are often too busy with their own projects to get involved in the process. With this in mind, the European initiative ETHNA System has designed an ethics governance framework that can be implemented following specific guidelines and tools that can help different types of institutions to promote and generate responsible research and innovation.

This multi-author book aims to present these practical contributions, their conceptualization and characteristics, as well as the experience gained from their application in different institutions. Its main objective is to provide a practical and useful guide that will help other institutions to start introducing Research Ethics effectively in their organizations. In this sense, the book clearly distinguishes itself from other existing publications in research ethics that have more theoretical and philosophical content, without providing practical guidance. In addition, and drawing on prominent international researchers, the book also includes topics of future relevance in the field that arise from the current push for (the realization of) an honest and open science.

The results of the book are an outcome of the ETHNA system project (developed in the years 2020–2023) with the participation of various members of the project as well as external researchers who stand out for their knowledge in the field of RRI and research ethics.

The book presents some of the key issues that institutions need to address when implementing RRI. More specifically, the book is structured in four sections that seek to define the theoretical dimension of RRI (“Foundation”), the experiences of implementation developed during the project (“Experiences and Lessons Learned”), the RRI tools (“Ethics Tools in Practice”) and the future main challenges on integrity, open access, gender, public commitment and artificial intelligence (“Looking into the Future: Main

Challenges"). The chapters are presented in a simple and straightforward language that has been discussed and reviewed by the Reviewing Committee.

These sections are briefly summarized below.

The "Foundations" Section introduces the theoretical basis of RRI. Thus, the first chapter focuses on clarifying the meanings when using the RRI framework within the general literature. More specifically the authors distinguish different models of RRI governance such as top-down and bottom-up models to further develop other options. The second chapter lays the theoretical foundations of the ETHNA System, explaining its principles and how it works. This chapter introduces the ethics participatory and deliberative nature of the implementation proposed by ETHNA as well as the components of the ETHNA System. A key aspect that is further elaborated in the third chapter is the role of stakeholders and how they can be engaged; it further addresses how their demands can be worked out or considered in the design, process and outcome of the research.

The "Experiences and Lessons Learned" section provides a general analysis of the implementation processes developed during the ETHNA project. So, its first chapter explains in detail the methodology and steps carried out during the implementation, highlighting the general barriers, drivers, and good practices observed in the six implementation processes completed during the project. The next three chapters delve into three specific implementation cases developed respectively at a non-governmental research institute, a research and innovation ecosystem, and a university. These three chapters focus on three particular cases, but it is hoped that they will be useful to other institutions developing an RRI framework. Despite the differences between the types of institutions, these chapters share a number of lessons from their experiences, observing common drivers and barriers.

The "Ethics Tools in Practice" section focuses on three tools that are central to the implementation of RRI, namely ethics committees, training and public engagement. As such, one chapter explains the meaning of the ethics committee and its role and relevance today. This chapter also provides some examples and references that are of interest to any implementer. Another chapter provides a systematic review of methods to promote RRI learning. Through an extensive review of the literature, the authors provide a broad overview of how to design and promote RRI learning. In addition, the chapter offers an analysis of the basic characteristics and steps to be taken when implementing such training. They also provide information on projects, occasions or methods that may be useful for any institution interested in promoting training on RRI. A third chapter tackles another central theme of RRI: public engagement. In this case, the analysis is linked to the experience of NewHoRRIzon, another Horizon 2020 project, showing a variety of activities and approaches needed to spur public engagement with research and innovation.

Finally, the last section of the book "Looking into the Future: Main Challenges" delves into some further challenges that research faces in key and diverse areas such as: scientific integrity, the promotion of open access, the promotion of gender equality, and the development of advanced technology systems. Each of these issues is examined in a specific chapter. The first chapter of the section is a text structured around the eight main contexts offered by the European Code of Conduct for Research Integrity, namely (1)

research environment, (2) training, supervision and mentoring, (3) research procedures, (4) safeguards, (5) data practices and management, (6) collaborative working, (7) publication and dissemination, and (8) reviewing, evaluating and editing. In this way, the chapter reviews the most current challenges in each point, reviewing empirical studies and guidelines that examine them. In turn, the promotion of gender equality in the R&I context is addressed in the following chapter. It critically analyses the two perspectives from which work is being done in the European context. One perspective advocates working on gender equality as one of the key areas of RRI, while the other argues for mainstreaming gender in order to integrate it into all dimensions of RRI. After an in-depth analysis, the chapter shows the tendency to use equality plans in organizations as the most effective mechanism for promoting effective equality between men and women in the research area. The following chapter presents the characteristics of Open Science (OS), its evolution and current challenges. It also provides a broad and clear conceptual framework, useful for engaging with current debates on OS. In addition, the chapter offers ideas on how OS can be developed in the near future. Finally, the last chapter analyses aspects related to the implementation of advanced technology systems, namely Artificial Intelligence-based ones, within Public Administration.

We hope that this book is of interest to a broad group of researchers and people involved in the administration of universities, research centers and research funding, as well as other stakeholders interested in responsible research and innovation. Overall, the various contributions offer a wide range of experiences and insights that are intended to be useful for institutions dealing with similar processes. In sum, we believe that this collective book can contribute to further advance the implementation of RRI and the promotion of an open science with and for society.

March 2023

Elsa González-Esteban
Ramón A. Feenstra
Luis M. Camarinha-Matos

Organization

Reviewing Committee

Ana Inês Oliveira	Uninova-CTS, NOVA University of Lisbon, Portugal
Angel Puyol	Universitat Autònoma de Barcelona, Spain
Carlota Carretero	Universitat Jaume I, Spain
Dietmar Lampert	ZSI Center for Social Innovation, Austria
Elsa González-Esteban	Universitat Jaume I, Spain
Filipa Ferrada	Uninova-CTS, NOVA University of Lisbon, Portugal
Gábor Szüdi	ZSI Center for Social Innovation, Austria
Laura Bernal	Universitat Jaume I, Spain
Lars Ursin	Norwegian University of Science and Technology, NTNU, Norway
Lisa Häberlein	European Network of Research Ethics Committees, Germany
Luis M. Camarinha-Matos	Uninova-CTS, NOVA University of Lisbon, Portugal
Marko Hajdinjak	Applied Research and Communications Fund, Bulgaria
Nicté García	Universitat Jaume I, Spain
Ramón A. Feenstra	Universitat Jaume I, Spain
Rosana Sanahuja	Universitat Jaume I, Spain

Technical Sponsor

ETHNA System. Ethics Governance System for RRI in Higher Education, Funding and Research Centres

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 872360.

Contents

Looking into the Future: Main Challenges

Foundations

Institutional Governance of Responsible Research and Innovation

Marit Hovdal Moan(✉), Lars Ursin, and Giovanni de Grandis

Norwegian University of Science and Technology, NTNU, Postboks 8900, NO-7491 Trondheim, Torgarden, Norway
marit.hovdal.moan@ntnu.no

Abstract. In this chapter, we analyse the debate around the implementation of responsible research and innovation (RRI) in Higher Education, Funding and Research Centres (HEFRCs). We will illustrate some proposals about how to implement RRI in HERFCs in a good way. Open and inclusive governance is key to fruitful implementation of RRI in these organizations. Governance in this context refers to ways of steering processes in a desirable direction, in this case in the direction of responsible research and innovation that is ethically acceptable and socially desirable. We will present and assess different models of governance and aim to provide ethical governance of research and innovation (R&I) inspired by the most convincing ideas emerged in the current debate.

Keywords: Ethics · Meta-governance · Citizen involvement · Retrospective governance · Prospective governance

1 Introduction

In this chapter[1], we analyse the implementation of responsible research and innovation (RRI) in Higher Education, Funding and Research Centres (HEFRCs) from an institutional governance perspective. Governance in this context refers to ways of steering processes in a desirable direction, in this case in the direction of responsible research and innovation. We present examples of RRI governance practices in a selection of HERFCs in Europe, which represent different modes of institutional governance, however they cover mostly top-down examples, as bottom-up experiences are less represented in the literature. Nevertheless, bottom-up governance is an ideal often voiced in theoretical discussions about RRI and Research and Innovation (R&I) governance.

We argue that these different modes of governance reflect different understandings of what it means to act responsibly in R&I, which correspond with two distinct conceptions of responsibility: a retrospective conception, according to which acting responsibly entails avoiding harm and correcting harm when committed, and a prospective conception, according to which acting responsibly entails contributing to doing good in the future. These two conceptions of responsibility, in turn, inform different narratives of

[1] The chapter is based on the ETHNA report *State of the art and best practices* [1].

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 3–18, 2023.
https://doi.org/10.1007/978-3-031-33177-0_1

what the main purpose of RRI-governance should be: making R&I better equipped to avoid future harm and thus *doing right; or* aligning R&I with the needs and expectations of society at large, thus contributing to *doing good.*

Drawing on the examples presented we suggest that bottom-up modes of governance seem especially fit to integrating principles of RRI in everyday R&I practices, when RRI is understood to entail *doing good.* Bottom-up modes of governance are prone to be open and inclusive and could thus be described as ethically desirable modes of RRI governance. However, the examples also indicate that some form of meta-governing structure is necessary to sustain bottom-up governance structures over time, which potentially can undermine their openness and inclusiveness.

2 Retrospective and Prospective R&I Governance

Two different conceptions of responsibility are reflected in different governance approaches to RRI. We refer to these as *retrospective* and *prospective* conceptions of responsibility. On a retrospective conception of responsibility, acting responsibly entails avoiding harm and correcting harm committed in the past. In this sense it has a "backward-looking" perspective on what acting responsibly in research and innovation entails in practice. By contrast, a prospective conception of responsibility, focuses attention on contributing to doing good in the future, thus taking a "forward-looking" perspective on the practical implications of acting responsibly in research and innovation [2]. In the current landscape of HEFRCs in Europe, some RRI governance practices reflect retrospective conceptions of responsibility, while others assume some version of prospective understandings of responsibility. This is a significant conceptual distinction in the analysis of RRI governance in HERFCs, since these two different conceptions of responsibility inform distinct narratives of what the purpose of RRI governance should be: making R&I better equipped to avoid future harm and thus *doing right* (this narrative assumes a retrospective conception of responsibility)*, or* aligning R&I with the needs and expectations of society at large, thus contributing to *doing* good (this narrative instead assumes a prospective conception of responsibility).

Retrospective notions of responsibility have traditionally translated into a governance of R&I practices concerned with avoiding harmful products or practices of science and innovation, with a consequent focus on risk governance. However, R&I governance processes "premised on formal risk-assessment, have done little to identify in advance many of the most profound [negative] impacts we have experienced through innovation" [3, 4]. Retrospective accounts of responsibility are inherently limited in guiding decisions related to the trajectories of R&I, both due to the narrow concepts of risk that they assume [3–5]. And the hierarchical, top-down, regulatory forms of governance that they seem to entail, which runs counter to the unpredictable, future looking, collective enterprise of science and innovation practices. In response to R&I governance models premised on retrospective conceptions of responsibility, "a number of multi-level, non-regulatory, forms of science and innovation governance models have taken [a] forward-looking view of responsibility (…) attempt[ing] to introduce broader ethical reflection into the scientific and innovation process" [3, 4].

2.1 Research Ethics as a Governance Baseline

A basic object of the ethical governance of research and innovation is research ethics, usually through the mandatory introduction of research ethics committees that oversee that research practices and goals are not causing harm or violating rights. As it often happens, this level of governance emerged in response to scandals and public outrage and were introduced both with an eye to prevent bad things to happen again and to restore trust in [6].

Mechanisms to ensure research ethics may look like typical examples of a retrospective view of responsibility, as well as having a narrow view of [7]. Yet, their mandate can be expanded to include elements of prospective responsibility. An example is provided by the Norwegian Research Ethics Act (2017) that gives the Norwegian higher education and research institutions a statutory responsibility for putting research ethics into practice in their organization. Most of these institutions have ethics committees in place, mandated to handle cases having to do with fraud and other forms of misconduct in research. Norwegian national research ethics guidelines define the recognized research ethics norms in which the higher education and research institutions have a responsibility to provide training. The guidelines are specific for disciplinary areas and are managed by corresponding research ethics committees: the National Committee for Research Ethics in Science and Technology (NENT), the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH); The National Committee for Medical and Health Research Ethics (NEM) [8].

The respective guidelines place a responsibility on the institutions to include the broader societal perspective in the research ethics assessments they make; a responsibility that is already assumed in their legal obligation to provide training and education in research ethics. However, despite the broader scope of the national ethics guidelines, it is not common for higher education and research institutions in Norway to take a more proactive, prospective responsibility for research ethics, which would assume a broader societal understanding of what research ethics means. The guidelines are not only directed at the institutional level but are intended also to promote such reflection and awareness in the researcher. This further ambition needs a more integrate effort to be realised. So, in spite of some expansive aspiration, this research ethics approach in practice remains retrospective.

At the University of Twente we find a retrospective ethics committee [1]. University of Twente made research ethics assessment required for all fields as of 2020. A discipline-specific system of research ethics committees has been established, consisting of four internal ethics committees: one for the social sciences, one for the engineering sciences, and one for the computer sciences. A central (fourth) committee was set up to monitor the three sectional committees. The ethics system is modelled on the recommendations in the SATORI project [9], which developed a standard for ethics committees. One important recommendation coming out of the SATORI-project was that of establishing discipline specific committees; another one was that of securing a degree of transdisciplinarity in the composition of the committees' members. The committees should thus have expertise in the area being assessed, as well as one from a neighbouring area, legal expertise, and should include a member from outside the organization.

The Universitat Jaume I of Castellón (UJI) in Spain also has an internal ethics committee system [1]. Here the ethical governance approach is blended with the aspiration to promote the social responsibility of the university, which suggests a wider ambition that tends towards a broad ethical governance. The former rector of UJI initiated a process to develop the UJI's social responsibility policy. A focus group was established with the mandate to develop a draft ethics code. The focus group consisted of university staff, students, and other stakeholders, including companies that the university collaborates with. The group drafted an ethics code that established the ethical values of the university, with a section on the Integrity and Responsible research practices. An ethics and social responsibility system was put in place to monitor and assess the implementation of the ethics code, including an ethics and social responsibility committee. The members of the Ethics and Social Responsibility Committee include staff, students, the general secretary of the university, the vice-rector of research, the director of UJI equality office, the director of the deontological committee (research integrity committee), as well as the ombudsperson for students. All issues related to breach of the Integrity and Responsible Research Practices Code are discussed in the committee. Moreover, the ethical assessment of the projects is carried out by a research deontological committee.

2.2 Retrospective Responsibility and Its Limits

In general, one characteristic of retrospective, or backward-looking, conceptions of responsibility is that they focus attention on one-off, time-limited acts, which are undertaken in the past, by identifiable agents, with adequate control and knowledge of the likely harmful consequences of the act (including unintended, yet reasonably foreseeable harm [10–12]. One problem with this focus on time-limited conduct in the pasts, is that processes, which were initiated in the past and are still ongoing, such as for instance research and innovation practices, and structures within which processes take place, such as the current, global academic incentive structure, fall outside the realm of evaluation when the question of responsibility for harm arises. Instead, backward-looking conceptions of responsibility is premised on an understanding of harmful acts as temporary deviations from a legal and social background structure that is assumed as normal [13]. The concern with time-limited harmful conduct thus also overlooks the fact that harm can be experienced, not merely as a one-off harmful incident, but as a persistent institutional reality, permeating everyday life, in a structural way. To accommodate forms of harm and injustice that are structural, and thus not timebound, we need a concept of responsibility understood as generated by "deeds already underway", to borrow a term from Hans Jonas [14] rather than as retrospectively generated by deeds already done. The target are patterns of action embedded in the cultural and material reality of a social group (cf. The notion of "structural violence" [15]). Such patterns produce harm or injustice even if no particular action can be singled out as wrong.

A second problem with retrospective conceptions of responsibility is that responsibility arises only if a harmful outcome can be linked to an identifiable wrongdoer. As Young explains, this 'identity condition', implies that one isolates "the one or ones liable (…) thereby distinguishing them from others, who by implication are not responsible" [13]. The identity condition is problematic also in the context of R&I practices, given the plurality of actors often involved in the knowledge production process, and the fact

that there is often no interaction between actors involved in the R&I process and those affected by the outcome. This is what Denis Thompson calls the "problem of many hands" [16].

Lastly, retrospective conceptions of responsibility only recognise harm that could reasonably have been foreseen. With respect to R&I processes, we do not always know what (harmful) effects in society they will have. Here risks, uncertainty and conflicting evaluations mirror the circumstances that have led some risk scholars to develop adaptive, ongoing and participatory strategies of risk management, for instance Klinke & Renn [16]. The backward-looking model of responsibility is ill-suited both to handle the dispersed agency and uncertainties of R&I, and to inspire practices that are responsible in the sense of enabling to manage unpredictable risks and unintended consequences.

Committee systems remain the dominant accountability and ethics assessment mechanism for R&I projects. The question is whether this is a suitable, and sufficient, governance mechanism for the purpose of integrating RRI in R&I processes. The Norwegian ethics committee system illustrates how an ethics committee that is originally built on the model of a retrospective, top-down committee system can be combined with more distributed governing mechanisms aimed at setting the rules of the game and encouraging and facilitating reflection, through disciplinary specific national guidelines, and the creation of temporary national fora for debate on issues of general interest, which raise ethical questions and dilemmas.

3 Prospective R&I Governance

A central premise underlying the concept of forward-looking obligations is that the responsibility to act so as to produce a desirable state of affairs, or to prevent bad outcomes in the future, increases proportionally with the capacity to influence others or our surroundings, be it peoples' rights and freedoms, society's basic institutions, or the environment or climate [14]. Science and technology have the potential to influence people, society and their environment in profound ways, in both a positive and a negative sense. Applying this understanding of a forward-looking conception of responsibility to R&I governance seems to entail at least two presumptions about the nature of science and the relation between science and society, both of which are debatable: (i) that potentially harmful trajectories of science and innovation can be identified and stopped or changed before new technologies are 'locked in' to societal practices and structures [18], and (ii) that the direction of science can be steered towards whatever society deems desirable. These objections, though, may in turn make some debatable assumptions. Objection (i) seems to rule out the possibility of what we may call a life-cycle, or adaptive management of the unintended consequences of science and technology. Of course, we cannot assume that such management is possible, but it is an option to be tested. Objection (ii) seems to assume that there is clear and majoritarian public opinion about what is desirable for society. So, while the objection raises an important point about the possibility of imposing directionality to technoscientific advances, it overlooks the fact that social opinions on what is desirable are constantly evolving and are renegotiated in light of new experiences, challenges and public debates and frames. So, again, a continuing, process view of steering seems more appropriate, if possible. Keeping in mind the limits of prospective aspirations, let us have a look at some examples.

In January 2020 the Norwegian Research Council's (NRC) new policy on open research came into effect. The policy addresses in a systematic, strategic way open research as well as RRI, and the involvement of stakeholders in R&I. The focus on these topics is not new to NRC; however, the policy is a first attempt at linking all the elements and integrating them into NRC's work in a systematic way, as part of NRC's new portfolio strategy [1]. Inclusion of stakeholders is fundamental to the way in which NRC works to realise the policy on open science. The involvement of stakeholders is an important part of NRC's new organisational strategy, involving among other things a shift to portfolio management; a development which is in line with the move towards mission-thinking and involvement at the level of the EU [19]. NRC also strongly encourages stakeholder involvement at project level as well.

Despite this commitment, the practice of stakeholder involvement remains difficult to realise and not easy to fit within the established working practices and constraints of R&I. A telling example comes from Digital Life Norway (DLN). DLN is a large Norwegian centre that promotes biotechnology research and innovation as well as transdisciplinarity. DLN has a prize for the "transdisciplinary publication of the year" open to publications authored or co-authored by researchers based in Norway. This provides a good observation point for stakeholder involvement, as it is a key feature of transdisciplinarity. Yet, very few publications reflect a significant involvement of stakeholders outside academia. In 2021 only one submission satisfied this criterion, and in 2022 none, so that the prize will not be awarded.

Another case is the Science Ombud at the University of Oslo. The idea of establishing a Science Ombud was to put in place a form of governance system that could monitor several issues around research integrity, broadly understood, and not limited to preventing fraudulent behaviour [1]. The Science Ombud has an advisory role and shall function as a low-threshold service for researchers employed. The cases that the Ombud handles are often about co-authorship (40% of the cases in 2019). But the mandate also includes issues related to other topics, although not as broad as the responsibility concept of RRI. The Ombud has no formal authority, and the idea is that researchers should be able to seek out a low-level independent body within the institution, to discuss and resolve what they themselves experience as ethically problematic issues. Confidentiality is an important principle in the functioning of the Ombud, both to ensure that the Ombud institution remains low-threshold, and that those who contact the Ombud do not'risk' anything. The Ombud can therefore not proceed with a case without the consent of the person who reports it.

It is worthwhile elaborating on what it entails in practice to introduce a forward-looking conception of responsibility as a guiding principle for R&I governance - in contrast to that of a retrospective one. Arguably it requires a fundamental shift of mindset towards acknowledging "the intrinsically normative aspects of science and technology, including risk" [5]. At the core of prospective conceptions of responsibility is the idea that assigning responsibility to an agent concerns "the forward determination of what is to be done", in order either to create a desirable outcome, or to prevent an undesirable one. The focus is not on a particular wrong committed by an identifiable agent who merits blame or punishment, but on "getting the right people and institutions to work together to producing a desirable outcome or preventing a bad one" [10, 14].

What matters for responsibility to be generated on the forward-looking model is the combination of an outcome that is deemed valuable (be it the prevention of a harmful outcome or the facilitation of a desirable one), and institutional capacity or power to affect whether the outcome is achieved or not. With respect to R&I governance, a prospective view of responsibility entails a shift in focus from "preoccupations with 'downstream' risk-governance" [5], to a broader interest in the governance of profoundly political, and therefore public, concerns about what kind of society we want - and do not want - and what kind of knowledge is required to get there. This raises a very thorny issue: who should set the agenda for research and innovation? A question that triggers the conflict between researchers' and innovators' freedom and social control over the object and goals of their research. Should researchers and innovators retain the autonomy of judgement that is often assumed to pertain to professionals with great expertise, or since society pays the bill, and bears the risks of research, the principle that "who pays the piper picks the tune" legitimately holds? Pressing this principle faces the additional problem that it is very difficult to stir from outside activities based on highly specialised knowledge. So, we need to look at the resources of governance.

4 Perspectives on Governance of R&I

Governance can be conceptualized as a distributed mode of governing involving other actors besides policy makers and the top-management. This allows "politics [to be] shaped through several and diverse initiatives and authorities" coming from … "networks and partnerships consisting of a range of public and private actors ([20] *our translation*). This conceptualization of governance emphasizes the bottom-up dynamic of governance and points to the fact that while "governance arrangements may be designed to serve a purpose, [they] can also emerge and become forceful when institutionalized" [21]. As Rip points out, there is an important analytical distinction to be made between the above conceptualization of governance understood as constituted by "bottom-up actions, strategies and interactions", on the one hand, and governance understood as a mode of governing that "opens […] up an earlier centralized arrangement and make[s] it more distributed, on the other" [21].

Landeweerd and colleagues [22] conceptualize governance in the R&I sector as "the set of processes by which it is taken that stewardship [i.e. management] over (…) science and technology practices (research, innovation, etc.) ought to be organized in continuous calibration with those practices. "This continuous calibration, or adjustment, must necessarily entail dialogue with those enacting science and technology practices, thereby allowing a range of actors, including "policy makers, researchers, industry and civil society groups and nongovernmental actors" to partake in the shaping of those practices. In this way, decision-making processes are embedded within practice itself, rather than centralizing the authority of decision at the policy makers level [22]. Landeweerd and colleagues definition of governance is an example of what Rip refers to as governance whereby previously centralized arrangements are made more distributed, in contrast to governance as bottom-up actions and interactions that may in turn become institutionalized. Importantly, the distributed authority that governance entails should not be confused with earlier self-regulatory governing regimes characterized by scientists governing themselves internally, based on codes of conduct [23].

The concept of governance expresses a shift in the discourse on how science should be regulated, from internal self-regulation by scientists based on codes of conduct, to external regulation, yet with the ambition of allowing the actors enacting science and technology a greater degree of autonomy and a voice in how the regulation is exercised. Governance is a non-hierarchical mode of governing, in the sense that it entails a move away from attempts at steering research and innovation towards predefined aims (expressed for instance in thematic funding programs), or by stable means, (such as economic incentives and predefined indicators of performance). Compared to old regulatory models of government, which articulate hierarchical co-ordination mechanisms based on [centralized] authority, the concept of governance expresses a mode of external regulation "that is more decentralized and open-ended" [3, 4]. Indeed, in contrast to government, "governance is distributed almost by definition" [21].

RRI literature describes various forms of steering research and innovation (R&I) in the direction of responsibility in a de-centralized, open-ended way. Kuhlmann and colleagues focus on anticipatory or tentative governance models [24], Rip and colleagues on "real-time and other forms of technology assessment" [25], Wynne on "upstream engagement" [26], and Van den Hoven and colleagues on "value-sensitive design" [27]. Others use the terms network- and interactive modes of governance to capture the essence of governance [28].

Guston's description of anticipatory governance practices at the Center for Nanotechnology in Society at Arizona State University (CNS-ASU) may serve as an example of what a multi-level, non-regulatory approach to steering R&I processes in the direction of responsibility entails in practice, with respect to governance tools [29]: "CNS-ASU unifies research programs… across three critical, component activities: foresight (of plausible future scenarios), integration (of social science and humanities research with nano-scale science and engineering), and engagement (of publics in deliberations). CNS-ASU also performs educational and training activities as well as public outreach and informal science education". Governance in the CNS-ASU case focuses on integrating reflexivity in research and innovation activities and coordinating meeting places between scientists from the natural and social sciences and lay citizens. It aims at influencing actors in networks not by top-down steering, but by coordinating and facilitating cooperation, leaving concrete aims of the R&I activity to the networks, and allowing for probing and failing in the process [24].

Echoing the case described by Guston, Strand and colleagues observe that "[t]he question of how to govern (…) R&I networks from the perspective of funding bodies and/or government (…) is rapidly transforming from policy perspectives based on central control and accountability to a perspective where coordination and stimulation are key concepts" [30]. Importantly though, governance is not purely about coordinating and facilitating, but may involve a mix of soft and hard(er) governing mechanisms. Hence, as Stilgoe et al. point out, the governance mechanisms of facilitation, coordination and stimulation are commonly complemented with more traditional "policy instruments such as normative codes of conduct, standards, certifications, and accreditations "[3, 4]. That said, the prerogative of de-centralizing authority contained in the concept of governance means that governance in the area of R&I denotes, as a minimum, the act of "open[ing] up science and innovation" [31] to a wider range of inputs. Some would argue that

this opening up entails creating new spaces of 'public dialogue'" [3, 4], which in turn seems to point to governance mechanisms that encourage and enable networking, broad inclusion and deliberation.

5 Why Involve Citizens in R&I Governance?

If to be responsible in R&I means to meet this ideal of a more representative co-construction, then responsibility entails democratizing research and innovation. Those affected by the new technologies in the future need to be involved in debating the shaping of that future, notably by participating in the framing of the problems and questions to be researched [32]. The focus here is on the process, where democratic procedures are thought to contribute among other things to "the awareness of a more local, historically and socially contingent knowledge production", and in this sense a more reflexive, "socially robust", knowledge and technology [33, 34]. Inclusion is an end, and not just a means to achieve a given end.

Importantly, as Randles and colleagues emphasize the demand for inclusion "is not just about inclusivity of a wider and more diverse range of perspectives, but that inclusion follows a co-construction ambition (…) [where] wider interests participate in the framing of research, innovation, and responsibility 'problems'; it is about how the processes of inclusion are constructed" [32]. A governance structure that aims at promoting and facilitating "upstream engagement" echoes the assumption that an inclusive, deliberative approach to science and innovation practices is an efficient mechanism for making R&I more reflexive, and - as a result - more anticipatory, and thus responsible.

The belief in the efficiency of upstream engagement as a mechanism to achieve more reflexive R&I practices has been justified with reference to the observation that "insight in the diversity of those participating in social-political interactions can only be gained by involving them in the governing process, considering them necessary sources of information" [35]. In a similar vein, Sykes and Macnaughten suggest that "choices concerning the nature and trajectory of [science and] innovation can be co-produced with publics in ways that authentically embody diverse sources of social knowledge, values and meanings" [36]. It has also been argued that research and innovation must engage with the public to serve the public [37, 38], and that "dialogue is the right thing to do for reasons of democracy, equity and justice" [36]. Others, however, have criticized the belief in public participation as an efficient mechanism for making R&I more reflexive, arguing that there is a lack of empirical evidence supporting its assumed quality and impact [39].

As pointed out by Landeweerd and colleagues. Above [22], responsibility in R&I is a matter of aligning science with the needs and expectations of society at large; that is, the goal of creating technologies that not only are not harmful, but also good, in the sense that they can be said to be socially, ethically, and environmentally desirable, and therefore also an expression of social priorities and informed preferences. If the main purpose of an R&I governance system is to ensure broad involvement in R&I processes, a relevant governance mechanism would be that of constructing good processes for involvement or rigging meeting places fit for that purpose; if, instead, the main purpose is to ensure that R&I contribute to solve the grand challenges of our time, a main governance mechanism

may rather be that of facilitating transdisciplinary collaboration, where involvement of lay citizens could be one element, but not necessarily so. Note that here we assume that the grand challenges are identified by experts. If the grand challenges were identified through public involvement, then this opposition would disappear.

6 Fine-Tuning Citizen Involvement in R&I

Public engagement governance tools have been criticized, among other reasons, for framing the participation exercises in ways that are useful to particular interests [40], for downplaying the low political status of the outputs of these exercises, and for serving as an "efficient tool of de-politicizing science and technology, in much the same way as ethics expert reviews" [22]. An ethics of involvement thus concerns not just the question of who should be involved in R&I processes and why, but the question of how the persons involved should be involved. This in turn, raises further questions: how those involved can participate on an equal footing with researchers, and how their contribution should be weighed in with that of researchers. These are questions that relate to the critique of public engagement exercises concerning the low political status of the outputs of these exercises. Furthermore, these questions raise the problem of how to weigh against each other etico-political and epistemological considerations, as well as how to protect the integrity of science. Science and technical expertise can be corrupted in different ways. They can be used to mask political choices under the pretext of techno-scientific requirements, but they can also be pushed to accept assumptions that do not meet their epistemic standards and incorporate value assumptions that are controversial and contested.

Landeweerd and colleagues [22] criticize the public participation model for taking a top-down regulatory form when put into practice, and for sharing the pitfalls of either frustrating the voice of "societal views and opinions or becom[ing] a scapegoat for pre-existing agendas". Landeweerd and colleagues argue that RRI as a mode of governance should link the governance of R&I to what von Schomberg has called "normative anchor points", such as sustainable development and social progress [41]. This move involves that the governance of R&I should no longer be restricted to "the definition and implementation of regulation in the form of negative constraints for science and technology but also of positive aims in a societal setting" [22], thereby broadening up the governance of science "to include topics and issues addressing community values and collective behavior" [22].

Moreover, the whole process of science - and not just its products – should be subject to transdisciplinary dialogue, meaning deliberation across disciplinary divides as well as with a variety of stakeholders, including the non-expert public. Acceptability and desirability assessments should thus take place from the outset of R&I processes, when problems are framed, rather than at the stage when a project is defined, or a product is ready to be introduced to the market. These assessments should take place at various stages throughout the process, and should involve a broad range of stakeholders, rather than being confined to scientific and ethical expertise.

RRI as a governance tool can be understood to move beyond the participatory governance approach "that merely emphasizes the inclusion of different actors", to designate

"the type of engagement that actors should exhibit in the process of doing research and innovation" in a responsible way [42]. The type of engagement that doing RRI entails can be summed up in the RRI dimensions articulated by Stilgoe and colleagues [3, 4]: anticipatory, reflexive, inclusive and responsive. Taken together, these criteria envision a continuous model of public engagement throughout the life-cycle of R&I. On Landeweerd and colleagues' account, RRI as a mode of governing entails opening up science and innovation in a way that allows for it being "shaped through several and diverse initiatives and authorities" through "a range of public and private actors" [20] (*our translation*). The move towards a governance of R&I activities can thus be understood as a response to RRI's normative commitment to opening up the shaping of science and innovation to society; to reduce – and even collapse – the society-science divide that informs, and is upheld by, the self-governing, technocratic and ethics expertise modes of governing R&I.

7 Meta-governance of R&I

We follow up on this by discussing different conceptions of ethical governance in HER-FCs. Our discussion takes us from top-down governing to bottom-up ideals of governance and their tensions, and further to the concept of *meta-governance*: facilitating the self-governance of networks through targeted procedural principles. These principles set the rules of the game and provide a common direction to R&I activities. Setting the rules of the game, however, is not a neutral intervention: it provides a frame and limits to self-governance. As the political theory of constitutionalism shows, procedures, frames and limits are ambivalent tools: they enable and they constrain, they confer power and they take away power. This is true for government as it is true for governance. Meta-governance sounds like a less intrusive concept than governance. But meta-governance is *the governance of self-governance.* As soon as we spell it out, the tension between intrusion and non-intrusion in the self-governance process becomes visible. We draw lessons and discuss the essential tensions emerging from the RRI literature on governance and meta-governance relevant for informing ETHNA System and similar RRI initiatives that aim to be open and inclusive.

The concept of RRI contains a dimension that designates responsibility as process, as well as a dimension that connects responsibility to particular outcomes [38]. Von Schomberg stresses that the process and product dimension of RRI are interrelated. The innovation process should thus be "responsive, adaptative, and integrated" and products developed through the innovation processes should "be evaluated and designed with a view to [the] normative anchor points [of environmental protection] (…) human health, sustainability, and societal desirability" [38].

Owen and colleagues [36] argue that a framework for what they refer to as "responsible innovation" must include consideration not only of the products of research and innovation, but more profoundly of the purposes and underlying motivations of R&I, by which they mean "not just what we do not want science and innovation to do, but what we do want them to do". This involves reflecting on "what sort of futures(s) we want science and technology to bring into the world, what futures we care about, what challenges we want to meet, what values these are anchored in" [36]. A core question here

is “how can the “right impacts” be democratically defined?” [36]. One possible answer to that question is by constructing a procedural framework that ensures fair deliberation on right impact.

Randles et. al argue [32] that the inherent normativity of RRI raises the question of “how to deal with the inevitable tensions, conflicts and related power games that arise when a heterogeneous, pluralistic actor landscape with diverging interests is confronted by norms and values intended to change behaviour”. Given the complexity of R&I networks that RRI as governance mechanism aims to facilitate, accommodate and strengthen (be it as a normative claim or a pragmatic move), the question is how best to deal with the inevitable conflicts and tensions that will arise in any “collective search for and foundation of normative direction” [32]. Randles and colleagues suggest that rather than contributing to this collective search for normative foundation, one should construct governance mechanisms “able to address contestation and facilitate the capacities and capabilities of the relevant actors to engage in constructive negotiations”, allowing the actors involved in R&I networks to negotiate the normative substance of the R&I activity themselves [32].

In a somewhat similar vein, Landeweerd and colleagues argue that “acknowledging complexity means that governance should be less about defining clear-cut solutions and more about making explicit the political issues that are at stake in science and technology. In this sense governance becomes a process in which the political nature of science and technology is made explicit, where concerned actors express that there is de facto not one, single answer (…) This means focusing less on decision-making and more on identifying the shared values and interests we have in the issues on the table; [the focus should be] on collaboration and dialogue, and on empowering participants” [22].

The RRI as governance approach on this procedural account “do[es] not focus on what RRI is (…) but on the processes and mechanisms by which it is thought to be realized” [43]; it is about providing an institutional framework that facilitates collective processes of cooperation, deliberation and negotiation, through a mixture of governance mechanisms. These include overarching principles for legitimate procedures and codes of conduct setting the rules of the game, the establishment of spaces for debate and negotiation, and policy instruments “helping to achieve legitimate agreements” [43]. Owen and colleagues [36] propose that a prospective conception of responsibility suggests an evaluative framework for what kind of processes qualify as legitimate in the governance of R&I, given the aim of steering R&I in the direction of responsible practices.

The ETHNA project [44] is a recent contribution to the RRI discourse on the governance of research and innovation (R&I). The proposed system of R&I governance in ETHNA includes four tools: an code of ethics and good practices in R&I, an ethics committee on R&I, an Ethics Line and, and indicators to monitor the progress and the performance [45]. The philosophical foundation of the ETHNA system - Habermas’s theory of communicative action [46] - presumes a procedural approach to governing research and innovation. The overarching aim is to steer R&I processes towards responsibility understood in a prospective, or forward-looking way. Governance theorists tend to agree that in order to enhance networks’ alignment with and contribution to a public good there is a need for “a system of meta-governance to stabilize key players’ orientations, expectations, and rules of conduct” [47–50].

As Jessop explains, "[m]eta-governance [is] the 'organization of self-organization'. It involves (…) the design of institutions and generation of visions which can facilitate not only self-organization in different fields but also the relative coherence of the diverse objectives, spatial and temporal horizons, actions, and outcomes of various self-organizing arrangements (…) [Organizations] have a major role here as the primary organizer of the dialogue among (policy) communities, as an institutional ensemble charged with ensuring some coherence among all subsystems, as the source of a regulatory order in and through which they can pursue their aims" [47]. The limits of such statements are the lack of specificity. One can make big claims about the virtues of meta-governance, but unless meta-governance is given a more specific content and it is tested in practice, it runs the risk of being a purely verbal, rhetorical solution. On the other hand, if meta-governance is specified into strict, pre-defined procedures and methods it runs the risk to be either context insensitive (and hence top-down) or not feasible within real-life settings (hence too abstract and ineffective). The ETHNA concept can be seen as an attempt to produce and test a prototype model of meta-governance.

The four principles of Owen and colleagues can provide a common RRI vision, and a common understanding of the rules of the game, in a given organization. As Sørensen argues, a meta-governance structure is needed to ensure that self-governing networks follow the rules of the game. If R&I networks are to contribute to solving societal grand challenges in a just and effective manner "they must be meta-governed with that purpose in mind", to paraphrase Sørensen [49].

The concept of a meta-governance structure succinctly captures the function that Owen and colleagues' four procedural principles can have in the governance of R&I in the direction of RRI, namely that of setting the 'rules of the game' and providing a common direction to R&I activities. In this sense the principles can be understood as constitutive of the regulatory order of R&I activities. The ETHNA system and similar systems of ethical governance of R&I can involve citizens based on a meta-structure in this sense. The four principles of Owen and colleagues could for instance inform the design and use of the four tools of the ETHNA System to involve citizens in the governance of R&I in a good way.

In the evaluation of the ETHNA system as an attempt to provide a concrete skeleton to the concept of meta-governance, it will be very important to test to what extent the ETHNA tools manage to negotiate the dialectic between, on the one hand, inviting participation and empowering bottom-up initiatives, and, on the other hand, offering a framework that ensures that such involvement and initiatives meet the values and normative principles of RRI.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission. The authors would like to thank the reviewers for their comments.

References

1. Hovdal Moan, M., et al.: ETHNA System: Literature review and state of the art description Mapping examples of good governance of research and innovation (R&I) related to responsible research and innovation (RRI), in Higher Education, Funding and Research Organisations (HEFRs) in Europe. www.ethasystem.eu and Zenodo (2022). https://doi.org/10.5281/zenodo.5814680
2. Owen, R., Macnaghten, P., Stilgoe, J.: Responsible research and innovation: from science in society to science for society, with society. Sci. Public Policy **39**(6), 751–760 (2012)
3. Stilgoe, J., Owen, R., Macnaghten, P.: Developing a framework for responsible innovation. Res. Policy **42**(9), 1568–1580 (2013)
4. Stilgoe, J., Owen, R., Macnaghten, P., Gorman, M., Fisher, E., Guston, D.: A framework for responsible innovation. Responsible Innov.: Manag. Responsible Emerg. Sci. Innov. Soc. **31**, 27–50 (2013)
5. Felt, U., et al.: Taking European knowledge society seriously: Report of the Expert Group on Science and Governance to the Science, Economy and Society Directorate, Directorate-General for Research, European Commission (2007). https://ec.europa.eu/research/science-society/document_library/pdf_06/european-knowledge-society_en.pdf
6. Pappworth, M.H.: Human guinea pigs--a history. BMJ: British Medical Journal **301**(6766), 1456–1460 (1990)
7. Carusi, A., De Grandis, G.: The ethical work that regulations will not do. Inf. Commun. Soc. **15**(1), 124–141 (2012)
8. FEK. https://www.forskningsetikk.no/en/about-us/ (2023)
9. SATORI. https://satoriproject.eu/ (2014–2017)
10. Cane, P.: Responsibility in Law and Morality. Bloomsbury Publishing, Oxford (2002)
11. Hart, H.L.A.: Punishment and Responsibility: Essays in the Philosophy of Law. Clarendon Press, Oxford (1968)
12. Honoré, T.: Responsibility and Fault. Bloomsbury Publishing, Oxford (1999)
13. Young, I.M.: Global Challenges: War, Self-Determination and Responsibility for Justice. Polity Press, Cambridge (2007)
14. Jonas, H.: The Imperative of Responsibility: In Search of an Ethics for the Technological Age. University of Chicago Press, Chicago (1984)
15. Farmer, P.E., Nizeye, B., Stulac, S., Keshavjee, S.: Structural violence and clinical medicine. PLoS Med. **3**(10), e449 (2006)
16. Thompson, D.F.: Restoring Responsibility: Ethics in Government, Business, and Healthcare. Cambridge University Press (2005)
17. Klinke, A., Renn, O.: Adaptive and integrative governance on risk and uncertainty. J. Risk Res. **15**(3), 273–292 (2012)
18. Arthur, W.B.: Competing technologies. Econ. J. **99**(394) (2017)
19. Mazzucato, M.: Governing Missions in the European Union. Independent Expert Report (2019)
20. Aars, J., Fimreite, A.L., Homme, A.: Styrt fragmentering og fragmentert styring–Nettverkspolitikk i norske kommuner. In: Rapport, Rokkansenteret (2004)
21. Rip, A.: The past and future of RRI. In: Rip, A. (Ed.): Futures of Science and Technology in Society, pp. 115–133. Springer, Wiesbaden. (2018). https://doi.org/10.1007/978-3-658-21754-9_7
22. Landeweerd, L., Townend, D., Mesman, J., Van Hoyweghen, I.: Reflections on different governance styles in regulating science: a contribution to 'responsible research and innovation.' Life Sci., Society Policy **11**(1), 1–22 (2015)

23. Tancoigne, E., Randles, S., Joly, P.B.: Evolution of a concept: a scientometric analysis of RRI. In: L. e. al. (Ed.), Navigating towards shared responsibility in Research and innovation: Approach, Process and Results of the Res-A-Gora project, pp. 38–44 (2016)
24. Kuhlmann, S., Stegmaier, P., Konrad, K.: The tentative governance of emerging science and technology—a conceptual introduction. Res. Policy **48**(5), 1091–1097 (2019)
25. Rip, A., Misa, T.J., Schot, J.: Managing technology in society: Pinter Publishers, London (1995)
26. Wynne, B.: Risk and environment as legitimatory discourses of technology: reflexivity inside out? Curr. Sociol. **50**(3), 459–477 (2002)
27. Van den Hoven, J., Lokhorst, G.-J., Van de Poel, I.: Engineering and the problem of moral overload. Sci. Eng. Ethics **18**(1), 143–155 (2012)
28. Kooiman, J.: Social-political governance: overview, reflections and design. Public Manage. Int. J. Res. Theory **1**(1), 67–92 (1999)
29. Guston, D.H.: The anticipatory governance of emerging technologies. Applied Sci. Convergence Technol. **19**(6), 432–441 (2010)
30. Strand, R.: Striving for reflexive science. Fteval Journal for Research and Technology Policy Evaluation (48), 56–61 (2019)
31. Stirling, A.: "Opening up" and "closing down" power, participation, and pluralism in the social appraisal of technology. Sci. Technol. Human Values **33**(2), 262–294 (2008)
32. Randels, S., et al.: Framing and frameworks: six grand narratives of *de facto RRI*. In: Ralf, L., et al. (eds.): Navigating towards shared responsibility in Research and innovation: Approach, Process and Results of the Res-A-Gora project, pp. 30–37 (2016). https://indd.adobe.com/view/eaeb695e-a212-4a34-aeba-b3d8a7a58acc
33. Nowotny, H.: The place of people in our knowledge. European Rev. **7**(2), 247–262 (1999)
34. Jasanoff, S.: Technologies of humility. Nature **450**(7166), 33 (2007)
35. Kooiman, J., Jentoft, S.: Meta-governance: values, norms and principles, and the making of hard choices. Public Admin. **87**(4), 818–836 (2009)
36. Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., Guston, D.: A framework for responsible innovation. Responsible Innov.: Managing Responsible Emergence of Sci. Innov. Society **31**, 27–50 (2013)
37. Braun, R., Griessler, E.: More democratic research and innovation. J. Sci. Commun. **17**(3), 1–7 (2018)
38. von Schomberg, R.: A vision of responsible research and innovation. In: Owen, J.B.R., Heintz, M. (eds.): Responsible Innovation : Managing the Responsible Emergence of Science and Innovation in Society, pp. 51–74. John Wiley & Sons Inc. (2013)
39. Rowe, G., Frewer, L.J.: Public participation methods: a framework for evaluation. Sci. Technol. Human Values **25**(1), 3–29 (2000)
40. Felt, U., Fochler, M., Müller, A., Strassnig, M.: Unruly ethics: on the difficulties of a bottom-up approach to ethics in the field of genomics. Public Underst. Sci. **18**(3), 354–371 (2009)
41. Von Schomberg, R.: Prospects for technology assessment in a framework of responsible research and innovation. In: Technikfolgen abschätzen lehren, pp. 39–61. Springer. (2012). https://doi.org/10.1007/978-3-531-93468-6_2
42. Burget, M., Bardone, E., Pedaste, M.: Definitions and conceptual dimensions of responsible research and innovation: a literature review. Sci. Eng. Ethics **23**(1), 1–19 (2017)
43. Lindner, R., et al.: Responsibility navigator: ResAGorA (2016)
44. ETHNA System Project (2020–2023). www.ethnasystem.eu
45. González-Esteban, E.: The ETHNA system and support tools. In: Ethics and Responsible Research and Innovation in Practice, Springer, Cham (2023)
46. Habermas, J.: Theorie des kommunikativen Handelns. Frankfurt, Suhrkamp (1981)
47. Jessop, B.: The rise of governance and the risks of failure: the case of economic development. Int. Soc. Sci. J. **50**(155), 29–45 (1998)

48. Jessop, B.: Governance and meta-governance in the face of complexity: on the roles of requisite variety, reflexive observation, and romantic irony in participatory governance. In: Participatory governance in multi-level context, pp. 33–58 (2002). Springer https://doi.org/10.1007/978-3-663-11005-7_2
49. Sørensen, E.: The metagovernance of public innovation in governance networks. In: Policy and Politics Conference in Bristol, pp. 16–17 (2014)
50. Sørensen, E., Torfing, J.: Theories of Democratic Network Governance. Springer (2016). https://doi.org/10.1057/9780230625006

The ETHNA System and Support Tools

Elsa González-Esteban(✉)

Universitat Jaume I, Avenida Sos Baynat, S/N, 12071 Castellón, Spain
esteban@uji.es

Abstract. The article deals with ethics governance systems in the field of research and innovation at the organisational level, both for organisations performing and funding research and innovation activities. In particular, it proposes and argues for a system called ETHNA System. Informed by a deliberative and participatory concept of ethics governance, as well as by the dimensions of responsibility in research and innovation – anticipation, inclusion, reflection and responsiveness – it proposes a modular design of ethics governance based on four mechanisms: a responsible research and innovation (RRI) Office(r); a Code of Ethics and Good Practices in research and innovation (R&I); an ethics committee on R&I and an ethics line. Moreover, to ensure continuous improvement, a system for monitoring the process and the achievement of results is provided. The system also offers specific details of the implementation process paying attention to four issues: research integrity, gender perspective, open access and public engagement.

Keywords: ethics governance · responsible research · university · RRI Office (r) · Code of Ethics and Good Practices · Ethics Committee · Ethical Line · monitoring indicators

1 Introduction

The ETHNA System (Ethics Governance System for Responsible Research and Innovation (RRI) in Higher Education, Funding and Research Centres) has been designed as a governance system that enables the ethical self-regulation of the research activity. It has been designed as a governance structure which should be ethical and effective at the organisational level.

The ETHNA System consists of a foundation block (an ETHNA Office or Officer) and three column blocks (or ethical tools) with a monitoring block. Together, these blocks align the spaces of research and innovation with the highest ethical standards of RRI.

Competences of the ETHNA System are linked to the R&I space of the university institution, promoting good practices aligned with the values, needs and expectations of their internal and external stakeholders. Its main responsibility is to promote ethical responsibility through self-regulation, while following the current legislative framework, but without resulting in a new bureaucratic compliance process [1].

The ETHNA System addresses a commitment that goes beyond professional self-regulation in the area of research ethics, demanding the self-regulation of the organisations that carry out or fund R&I so that they live up to ethical standards and to society's

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 19–37, 2023.
https://doi.org/10.1007/978-3-031-33177-0_2

needs and expectations while respecting the existing legislative framework. As several authors point out, "self-regulation and behavioural norms within research communities alone have not prevented all sorts of fraudulent behaviour in research" [2] so the presence of structures in these organisations that will ensure a culture and environment inclined towards ethical excellence is necessary in the 21st century. In Chapter 1 of this book [3], it has already been argued that this system follows a meta-governance model that implies the governance of self-regulation.

As will be seen in the following sections, this proposal for the governance of self-regulation of processes and practices in Research Performing Organisations (RPO's) and Research Funding Organisations (RFO's) is informed by a deliberative and participatory concept of responsibility and the governance of research. In morally pluralistic societies it is necessary to discover what values, principles and good practices could be accepted as correct or just, as there are no closed lists or unique codes to attend to. On the one hand, in these morally pluralistic contexts, discursive processes are required to facilitate the inclusion of those affected (in the present and the future) when determining the shared values and principles. On the other hand, given the complex organisations, as well as deliberating when there are conflicts of interest [4, 5]. The proposed system and its accompanying elements have been designed using this co-creation methodology, based on living-labs that have involved the participation of internal and external stakeholders (see [6–9]).

2 The "Ethics Governance" Concept that Informs the ETHNA System

Science and Technology Studies (STS) have given a good account of the shift in governance from self-regulation to external regulation in the late 1970s and 1980s. This includes a shift of who advises on governance: "First, those with scientific, technological expertise, with their input being guided by legal expertise; second, those with 'ethic-legal' expertise; and third, those included as part of 'public participation'" [10]. At that time, three styles of science and technology governance were present.

In the "technocratic" style of governance, a specific format for decision-making is dominant. This style implies two aspects of technical regulations: scientists and technologists as assessors of acceptable risk; law and lawyers as framers of governance procedures in order to, for example, make suggestions for changes to legal frameworks, self-regulation or new regulations.

In the "applied ethics" style of governance, the ensuing approaches were based on input from ethical experts from the fields of applied ethics and bioethics, as well as socially engaged scientists with specific experience and interest in ethical issues related to science and technology. Ethics in the governance of science and technology arguably has an influence in an advisory capacity on the moral issues that are intrinsically connected to science and technology, but also as a mediator regarding the outline of debate platforms in terms of the triple helix of transparency, democracy and trust.

In the "public participation" style of governance, from the late 1970s onwards, more proactive approaches to governance were developed to directly involve citizens in

decision-making on science and technology, be it for surveying public opinion, consultation, or direct democratic decision-making. These included citizen juries, citizen panels, consensus conferences, planning cells, deliberative polling, focus groups, consensus building exercises, surveys, public hearings, open houses, citizen advisory committees, community planning and referenda [10].

Landeweerd et al. [10] argued that RRI fitted in with the idea of moving from "governing" to "governance"; rather than locating the authority of decision at the policy-making level, governance aimed to embed decision-making processes into practice itself. Where the public and or citizen participation is the key.

The foreseen perils were that this framework would potentially damage not only the autonomy of the expert communities involved, but also the sovereignty of the public bodies (politicians, policy-makers) that ought to guarantee the public legitimacy of the choices made. In this case, although cooperation between public and private may seem to enhance the embedding of R&I in society, it may actually render public funding and public interest sub-service to private interests. Increasing the extent to which these interests serve both public and private goals may be a positive development, but does not mean that a voice for public interest is no longer needed. According to Landeweerd and colleagues [10], RRI can only be successful if it develops strategies to prevent publicly delegated sovereignty from eroding.

There are several implications to consider in an ethics governance system of RRI in the RPO's and RFO's. First, approaches to governance need to move beyond the idea of governance as "quick fixes" to ethical issues of science and technology. It needs to be acknowledged that nothing is clear-cut or well-defined. Not only is there a complexity of problems and uncertainty scenarios, but also a kaleidoscope of ethical, social, legal and technical issues requiring reflection and multidisciplinary work. Second, "acknowledging complexity means that governance should be less about defining clear-cut solutions and more about making explicit the political issues that are at stake in science and technology" along with the ethical issues that require inclusion and open deliberation. Third, new more hybrid styles of governance emerge, in which the role of expert knowledge is explicitly acknowledged, but the range of relevant forms of expertise broadens [10].

2.1 Toward Discursive Ethics Governance

The governance of science and technology is dominated by a risk-safety-and precaution discourse. Responsibility in governance has historically been concerned with "products" of science and innovation, in particular impacts that later become unacceptable or harmful to society or the environment. This approach doesn't encourage debate and reflection on frameworks for justice, welfare standards for marginalised groups, politics of exclusion, privacy, etc. Its responsibility approach is retrospective, to use the terminology used in [3, 10, 11].

This ethical paradigm suggests that conceptions of responsibility should build on the understanding that science and technology are not only technical, but also social, political and ethical. In this vein, governance processes also imply identifying "in advance many of the most profound impacts that we have experienced through innovation" [11]. Debate and controversies often surround science and technology and should be involved in governance. As Stilgoe and colleagues [11] argue "Such controversies have demonstrated

that public concerns cannot be reduced to questions of risk, but rather encompass a range of concerns relating to the purposes and motivations of research, joining a stream of policy debate about the directions of innovation. Yet, despite efforts at enlarging participation, current forms of regulatory governance offer little scope for broad ethical reflection on the purposes of science or innovation" [12]. This paradigm thus points to a prospective or forward-looking conception of responsibility where procedural ethics plays a relevant role.

The literature review argues that the consequentialist model of the ethics governance of research and innovation should move toward another model. Dissatisfaction with "risk-based regulation has moved attention away from accountability, liability and evidence toward those future-oriented dimensions of responsibility – care and responsiveness – that offer greater potential to accommodate uncertainty and allow reflection on purposes and values" [11, 13, 14].

The discourse that emerged from the European Commission (EC) at the start of the past decade used this ethics governance to focus on aligning R&I to the values, needs and expectations of society, and to move toward an ethical paradigm of the governance of science and innovation. It is an approach that is making great progress but still evolving, in what is now being called Open Science, but maintaining the requirements of RRI [15, 16].

In this approach, the EC attempts to strongly emphasise "societal grand challenges" that have been defined as six keys or the policy agenda. The problem today is, as many authors show in the literature on RRI [17], that for the EC these keys are used as the RRI framework and not as challenges that the RRI framework should tackle.

The ETHNA System considers that the ethics governance of the RRI model should be designed in the RPO's and RFO's, such as universities, technological parks, innovation centres, etc., to provide an answer to the four dimensions of the RRI process (anticipation, inclusion, reflection and responsiveness) in which the six keys or the policy agenda will be issues or societal grand challenges. These issues or keys do not constitute all the aspects to be addressed by a RPO or RFO, but the political agenda has identified them as relevant demands and social requirements for which knowledge has begun to be mobilised and the European research space has begun to be transformed around. RPO's and RFO's must therefore consider themselves open to other topics that may be relevant in their social or community environment. When faced with these, research centres should institutionalise anticipation, inclusion, reflection, and action processes.

2.2 Discourse Ethical Paradigm: RRI as a New Social Contract

RRI is the reflection of a new social contract between science, innovation and society in general [18] which involves a change in the division of moral labour in R&I [19] aimed primarily at research governance [20]. As Stahl remarks, "RRI is concerned with creating a new mode of research governance that can transform existing processes with a view to ensuring a greater acceptability and even desirability of novel research and innovation outcomes" [21]. Von Schomberg stresses that the product and process dimensions are naturally interrelated. Where "products should be evaluated and designed with a view to these normative anchor points: with a high level of protection to the environment and human health, sustainability, and societal desirability" and where the challenge of

the process dimension "is to arrive at a more responsive, adaptative, and integrated management of the innovation process" [14].

Until the mid-20th century, the social contract implied "freedom, social licence and funding to invent, innovate and pursue scientific endeavours" that have been exchanged for the promise, and sometimes expectation, of knowledge, understanding, and value (economic, social or otherwise). By including the expectations that existing norms, laws and standards of conduct are adhered to, there is a long-standing history of responsibility in the research integrity context in this regard. This contract has been re-evaluated, especially for the often unintended and unforeseen impacts [21], and to form complex interactions with, and transformative consequences for, society. This is a symptom of what Hans Jonas described as the "altered nature of human action", mediated through technology and innovation [13].

2.3 Models or Mechanisms to be Used to Manage this New Social Contract

The literature categorises them as two main models: *Old models of governance* characterised by being centralised and having a regulatory centre; *modern models* whose central feature is that they are decentralised, with open-ended governance and soft governance, operating in new places: markets, networks and partnerships, as well as conventional policy and politics [11, 22]. These new models include multi-level, non-regulatory forms of science and innovation governance combined with old models. There is a combination and complementation with policy instruments such as normative codes of conduct, standards, certifications and accreditations that run alongside expert reports, technology assessments and strategic roadmaps. They have attempted to open up science and innovation to a wider range of inputs, notably by creating new spaces of "public dialogue" [12, 23, 24].

As Stilgoe et al. explain [11], the conventional model focuses on technological product questions meanwhile ethics governance and research integrity move into processes and autonomy and well-being of human volunteers and animals involved in experimentation. The new social contract is in the uncertainty scope. It moves in a scenario where the answers are not given, they are open and have to be defined.

The limitations of the old models of governance to responsibility is evident especially if we focus on R&I because regulatory and governance forms do not exist in this area. Then the key question is how to build a *democratic governance* mechanism of the intention without biasing the innovation process at the beginning of the procedure, in line with the previous section.

The answer is an ethics governance of RRI that ought to be democratically inclusive. That is to say, RRI that will reflect on its intentions and will be aware of the future that we want and that concerns us, whose challenges we wish to face. And this kind of RRI should respond to a mixed ethics governance model where the value-oriented and culture approach of self-governance will take pre-eminence over the compliance or regulatory approach [25]. The implementation of RRI activities through research governance measures can employ widely *shared principles of research governance*, such as the *integration of democratic principles* into research, the *precautionary principle*, the *principle of regulatory parsimony* [20] and the *discourse ethics principle* [26].

In our understanding, the ETHNA System is a suitable option to shift research performance towards ethics governance using the modern governance style. Following this ethics governance relevant organisations will adopt RRI as a component of their strategic framework and will aim to ensure all R&I activities cover all (or most) RRI components: the dimensions and key issues or areas. From this point of view, ethics governance of RRI is not just about compliance and a centralised model. It is about a set of aspirations related to a continuing commitment to be proactive, inclusive, reflexive and responsive with the ethical challenges that R&I has to face. Then ethics governance of RRI in RPO's and RFO's should pre-eminently be a deliberately decentralised self-governance model.

3 A Model for the Ethical Self-regulation of Research for Universities Based on the RRI Framework: Generating Open Science

The key issue is addressing the ethical concept lying beneath RRI and its institutionalisation through a government system. It is an ethical concept that needs a critical meaning, and not merely a conventional one. We will then be able to take account of both backward-looking and forward-looking responsibility [27]. It is also necessary to design ethical management systems of RRI dimensions (anticipation, inclusion, reflection, responsiveness) in institutions and organisations that can provide answers in key areas (integrity research, gender perspective, open access, public engagement, among others) when organisations generate or fund R&I [28, 29].

RRI emerged in Europe and was driven by European institutions. The ELSA (Ethical, Legal and Social Aspects of Emerging Sciences and Technologies) framework was the forerunner of the current concept. Initially, the aim was to lay down some guidelines in the research used as regulatory frameworks to be considered in developing European R&I [17, 19, 30, 31].

In 2011, in an expert group context, the European Union (EU) presented the RRI concept, defined in 2012 as follows: "Responsible Research and Innovation means that societal actors work together during the whole research and innovation process in order to better align both the process and its outcomes, with the values, needs and expectations of European society. RRI is an ambitious challenge for the creation of a Research and Innovation policy driven by the needs of society and engaging all societal actors via inclusive participatory approaches" [32].

This definition catches up a new approach in the RRI concept driven by the EU since 2011, that has been an "anticipatory governance", a guideline for R&I activity. Anticipatory governance aims to create a research atmosphere in which all present and future scenarios that can help to minimise possible risks to decision-making by providing feasible alternatives can be highlighted. Moreover, the integration of social science and humanist perspectives is important as a control, and also because it creates opportunities for dialogue and more reflexive decision-making [33, 34]. Additionally, this "anticipatory governance" entails a "democratic governance" that promotes interaction among the several agents integrating heterogeneous values, concerns, intentions and purposes [35]. The underlying idea is that research and innovation need to be democratised and must engage with the public to serve the public [14, 36].

RRI becomes an EU requirement so that the scientific community and society can work together to make the processes and results of science respond not only to the expectations, values and reflection of researchers, but also to those of citizens [29]. RRI can therefore be claimed to be a concept that comes from EU scientific legislators and institutions in a top-down process [29, 37]. However, at the same time, the RRI concept and its implied practice are also a bottom-up process in which existing experiences should be taken into account, as well as encouraging mutual learning [38].

It should be noted that some studies show how researchers and the scientific community recognise RRI traits and identify responsible R&I features with the same relevant issues [39]. Yet, at the same time, these scientists affirm that some practical barriers hinder RRI which could be tackled with a range of strategies at different levels. Recent studies have shown that: "These constraints included barriers likened to time, funding, reward systems, training, expectations of scientific production and the moral division of labour" where some are at an individual level but others are institutional [39].

As noted, the ETHNA System has been designed to be a way how RRI practices can be taken root in organisations that fund or carry out R&I. The aim is to drive ethical R&I using the double loop of "top-down" and "bottom-down" processes and, in parallel, as a means to overcome some obstacles for RRI and to promote it. The ETHNA System attempts to do so by working at an institutional level to promote RRI at an individual level, as both levels are interrelated or intertwined.

3.1 Three Complementary Discourses Behind the RRI Framework

The RRI concept then arises as a path to three discourses: democratic governance, responsiveness and responsibility that arises [12, 15, 16, 40]. The first discourse emphasises the *democratic governance* of R&I purposes and their orientation towards the "right impacts". Discourse on democratic governance is linked to reflection on the purposes and motivations for the products of science and innovation made in accordance with democratic governance. The RRI discourse asks "how the targets for innovation can be identified in an ethical, inclusive, democratic and equitable manner" [40]. Its purpose is therefore to democratically open up and materialise new areas of public value for science and innovation where the principle of participation plays an important role. The early definition shows the need to share values in which R&I are anchored. Then the next question is "What are the 'right impacts' of R&I?", and what values should these be anchored to?" [41, 42]. As previously mentioned, there are several proposals as to where the correct impacts can be defined (European Treaty, Human Rights Declaration, SDG among others), as well as different methodologies to be able to recognise them.

The second discourse focuses on *responsiveness* by emphasising the integration and institutionalisation of established approaches of *anticipation, reflection and deliberation* in and around R&I, influencing their direction and the associated policies.

Discourse on responsiveness involves reflection on R&I consequences and implications, intended and unintended impacts – anticipation, reflection and inclusive deliberation – on policy- and decision-making processes. It leads to the integration and institutionalisation of established mechanisms of reflection, anticipation and inclusive deliberation in and around R&I processes. And this reflection should be made on underlying purposes, motivations and potential impacts, what is known and what is not known,

associated uncertainties, risks, areas of ignorance, assumptions, questions and (ethical) dilemmas. There is a need to inclusively *open up* such reflection to broad, collective deliberation through processes of dialogue, engagement and debate by inviting and listening to wider perspectives from public and diverse stakeholders. Thus, RRI is concerned with the democratic governance of intent.

The third area concerns the framing of *responsibility* itself within the R&I context as collective activities with *uncertain and unpredictable consequences*. There are different subjects with responsibility: scientists, universities, innovators, businesses, policymakers and research funders. Funders play a leadership role in establishing a framework for RRI, but they must also lead by example. *Consequentialist* models of responsibility are *problematic* when we analyse the responsibility of innovation. In this context, other categories have been proposed as being more appropriate, such as *care* or *responsiveness* [18, 40], the definition of positive benefits or right impacts [14] and the legitimate interests of those affected [26, 43].

3.2 Procedural Ethics Governance with Four Dimensions for Approaching Ethical and Social Demands

In this section we look at the transition from theoretical to practical frameworks. That is, from the why to the how with a RRI operationalized. The ETHNA System has been designed to generate organisational processes that encourage responsible R&I following the RRI framework proposed by Owen, Stahl and Stilgoe, thus "entailing an ongoing commitment to be anticipatory, reflective, inclusively deliberative and responsive" [18]. Based on these institutionalised processes in RPO's and RFO's and using the building blocks, it is proposed to create the agenda of issues to be addressed and to which a response must be given in order to meet society's demands and requirements and to involve it in the very process of creating research and innovation. Specifically, the ETHNA System proposes the topics of integrity, gender, open access and public engagement, because it is necessary to begin by thematising aspects to be covered within the ethics governance system. And these issues have received much more attention in the European Union's research space, marked by the agendas of scientific policy and society. But, as has been pointed out, these issues were initially considered but there is room for identifying and dealing with further issues also based on stakeholder dialogue.

These four dimensions are not entirely new as they have their own background and roots. However, what is new in the RRI framework is that they can be used theoretically and practically as a procedural framework to guide R&I under conditions of uncertainty and ignorance. Moreover, "this redrawing will need to be done in a way that allows the constructive and democratic stewardship of science and innovation in the face of uncertainty toward futures we agree are both acceptable and desirable: this is a collective responsibility" [18]. The four dimensions of RRI could be described as a methodological process that implies their interrelation. The following is a brief account of these four dimensions and how they are covered by the ETHNA System [11, 18].

Anticipation [anticipatory]. This focuses on describing and analysing the intended and potentially unintended impacts that might arise whether they be economic, social, environmental, or otherwise. It is supported by methodologies that include foresight, technology assessment and scenario development, among others. The anticipation method

has been used to design the ETHNA System itself, holding panels with experts concerning both ethical and RRI governance, as well as stakeholders from inside and outside the research practice. It is also proposed as one of the processes to be strengthened within good practices in research, within the Code of Ethics and Good Practices.

Inclusion [deliberative and engaging]. This aims to inclusively open up visions, purposes, questions and dilemmas to broad collective deliberation through processes of dialogue, engagement and debate by inviting and listening to wider perspectives from public and diverse stakeholders. This is a part of a search for legitimacy and for generating trust in science [44]. This could include small-group processes for public dialogue: consensus conferences, citizens' juries, deliberative mapping, deliberative polling and focus groups. Inclusion and deliberation are argued as a "moral obligation" to identify the desirable outcomes of science and technology for society [45], and also as a way to accommodate new claims in discourse [46, 47]. Here we find important critics on effectiveness and its benefits. While there has been resistance shown to attempts to proceduralise public dialogue for fear that it becomes another means of closure or technocracy, efforts have been made to develop criteria that aim to assess the quality of dialogue as a learning exercise (see [6]).

Reflection [reflective]. This consists of reflecting on underlying purposes, motivations and potential impacts; what is known (including areas of regulation, ethical review, or other forms of governance that may exist) and what is not known; associated uncertainties, risks, areas of ignorance, assumptions, questions and dilemmas. There is a demonstrated need for *institutional reflection* in governance [26, 48]. The ETHNA System endorses four mechanisms designed to promote reflection through dialogue: the Code of Ethics and Good Practices in R&I, Ethics Committee on R&I and the Ethics Line. With this structure, RPO's and RFO's are promoting the adoption of standards, and may build this second-order reflection by drawing connections between external value systems and scientific practice [49].

Responsiveness [responsive or active]. This should be an iterative, inclusive and open process of adaptive learning with a dynamic capability. Responsible innovation requires the capacity to change shape or direction in response to stakeholder and public values, and changing circumstances. In order to generate ethics governance of research in RPOs and RFOs that respond to this dimension, the ETHNA System proposes to monitor the construction of governance structures (RRI Office(r), Code of Ethics and Good Practices, ethics committee on R&I, ethics line) with indicators of progress and performance. With its structures, it also establishes the processes for responding to the expectations and values that a RPO or RFO is expected to meet. On the one hand, it means to *respond* and, on the other hand, it means *to react and to answer.*

The ETHNA System is coherent with the procedural dimensions described [50]. Achieving this depends on scientific awareness, as well as on continuous and iterative processes that require the institutionalisation of governance [17].

3.3 Building an Open Agenda of Topics

The ETHNA System has been designed so that each RPO or RFO can address the issues that it considers an institutional priority. However, as a result of the collaboration within the European-wide Horizon 2020 project, not only has the design of the Ethics Governance processes been addressed, the aspects that could or should be covered have also been considered based on the four key points considered central by the RRI frameworks promoted by the EC. These are integrity research, gender perspective, open access and public engagement.

With its support tools, the ETHNA System offers an entry point to these issues that can be helpful for beginning deliberation and internal and external participation on how to define such issues, as well as thematising the demands that internal and external stakeholders are making towards RPO's and RFO's in order to trust their research, both as a process and as a result.

The outcomes regarding the issues are based on research carried out from 2020 to 2022 as part of the ETHNA System Project. This study involved a review of the scientific literature and previous European projects focused on RRI in different types of RPO's and RFO's, 23 in-depth interviews at European level with RRI and ethics governance experts, and a consultation based on semi-structured interviews and international workshops on different themes, as well as an online European survey covering internal and external stakeholders following the quadruple helix model [51]. The conceptualisation of these four issues after the study was defined as follows.

Research integrity is understood in the terms expressed by the European Code of Conduct for Research Integrity [52], which notes integrity in a series of basic principles, such as honesty, reliability, responsibility and respect. These principles concern the entire research process, from the initial approach to its execution to its dissemination, bearing in mind that quality research cannot take place if it does not meet scientific integrity criteria. Aspects that were shown to be important included authorship in research (publications and patents), originality, peer review, conflicts of interest and collaborative working.

The *gender perspective* in research is understood to promote the gender perspective at organisational level and at researcher and individual level in terms of any policies the organisation may have to promote equality and potential good professional practices among its members. Without attempting an exhaustive list, the ETHNA System identifies some core issues that should help institutions improve governance in terms of effective equality between men and women, as well as aspects such encouraging the equal participation of men and women in research teams at all levels; including experts sensitive to gender balance in the assessment process for R&I funding projects; and rewarding gender-sensitive research and innovation in the configuration of teams when applying for funding.

Open access in research and innovation is defined based on the policies and good practices an organisation can develop in terms of research results, data management, administration, management of intellectual property rights, and patents. Here, hot topics include data management and administration, storage and preservation, and results protection management.

Finally, *public engagement* shows commitment to citizen science and to science open to participation and deliberation from its audiences from the beginning of the

research design to the end, including development. In this sense, the essential issues are the processes for involving internal and external stakeholders particularly in research, taking into account the specific nature of vulnerable groups that may find it difficult to have a voice in research processes.

4 Description of the ETHNA System and Its Support Tools

This section presents a summary of the ETHNA System. First, it summarises the compass criteria that guide it and provide a basis for it, as well as the managerial principles that should guide its implementation in a RPO or RFO. Second, it shows that we are looking at a flexible system that starts from the material and leadership resources available to the organisation. Thus, the result of the implementation of the ethics governance system will take different forms, depending on how robust these two variables are in the organisation. Third, the three levels of commitment to which the organisation can adhere are presented. The process that backs the implementation of these levels is accompanied by support tools provided by the ETHNA System. Finally, the building blocks or basic ethics governance structures offered by the system are briefly described. Its implementation is understood based on the philosophy of continuous improvement.

4.1 Compass Criteria and Core Implementation Principles

The ETHNA System has two normative compass criteria that guide it and at the same time support it, and six managerial or pragmatic implementation criteria.

These two compass criteria guiding the ETHNA System are ethical and effective criteria. The ethical criterion, informed by Habermas' theory of communicative action [53–55], provides a compass to qualify the governance structure, based on the "all-affected principle", as more or less just. From the ethical foundation of this discourse, an ethics governance system is defined as one that promotes and facilitates (i) the inclusion of those immediately affected by it (i.e. R&I actors) in processes of discursive justification of the way in which the governance system is organised, and (ii) the inclusion of stakeholders (citizens, end-users, non-governmental organisations, business representatives, policy-makers) in processes of critical examination and discursive justification of possible scenarios and potential impacts generated by research and innovation processes. As can be seen, the ethical criterion is normatively based on the ETHNA System approach and, critically, on the four dimensions of RRI (anticipation, inclusion, reflection and responsiveness).

Moreover, the effective criteria, informed by governance theory on public innovation [56–58], refers to one that accommodates and facilitates the form that R&I activities often take, namely the form of networks. The networks are deliberative when defining the goals and objectives, highly autonomous in their working purpose and highly dynamic in their work processes.

The uptake of the flexible ethics governance ETHNA System compass by these two compass criteria needs core principles to help the implementation process. The core implementation principles that have been informed by the literature review and the process of interviews with experts that took place during the year 2020 following the "de

facto governance" view are [19, 59]: flexibility, adaptability, integrity, responsiveness and proactivity, networking and directness. Briefly, these principles can be understood as follows in the ETHNA System:

- Flexibility: the ETHNA System establishes a structure, protocol and an entry point and the organisation decides the level of commitment it wants to make to the ETHNA System.
- Adaptability (evolvability): the ETHNA System provides a guide to define and engage the organisation's stakeholders related to the R&I activities. The guide considers the plurality and diversity of the R&I ecosystem but not all the scenarios, so the ETHNA System should be adapted to each RPO or RFO.
- Integrativity: the ETHNA System promotes the integration of the R&I initiatives, procedures and structures running in the organisation related to the RRI dimensions and keys or issues.
- Responsiveness and proactiveness: the ETHNA System encourages the avoidance of bad practices or misconduct in R&I practices and relationships and promotes good practices.
- Networking: the ETHNA System endorses networking following the Quadruple Helix Method (QHM) and sustained collaboration with its internal and external stakeholders.
- Directness: the ETHNA System seeks to be direct and reduce the complexity a self-regulation system taking into account stakeholders and using the QHM could have.

4.2 An ETHNA System to be Built

As stated, the ETHNA System is a flexible ethics governance system that enables ethical self-regulation of the research activity at RPO's or RFO's. It has been designed as a meta-governance structure that is ethical and effective at an organisational level.

The ETHNA System offers four building blocks that incorporate a process for monitoring implementation and performance progress: a foundation block (RRI Office(r)) and three column blocks (Code of Ethics and Good Practices, Ethics Committee on R&I, Ethics Line). Together, these blocks align the research and innovation spaces with the highest ethical standards of RRI because they are designed to allow for anticipation, inclusion, reflection, and responsiveness in the institution's R&I space.

The procedural and formal structure of ETHNA System promotes the alignment of resources and processes already existing in the institutions, as well as the institutionalisation of new resources and processes, in order to implement a system of ethics governance of research (Fig. 1).

This shows a figure consisting of the different components of the ethics governance system (ETHNA System). Elements of the figure can be chosen and implemented in a flexible self-governance by RPO's and RFO's.

To further increase its adaptability, the ETHNA System regards two relevant factors as essential for the institutionalisation of RRI: the *leadership*, including the support it provides, on the one hand and the *base* on the other hand, i.e., the organisation's research staff with their values, awareness, skills, knowledge, and practices already in place. The latter may vary, depending on the organisational unit and research or innovation field. Both axes need to become strong in the long run. For further information on those

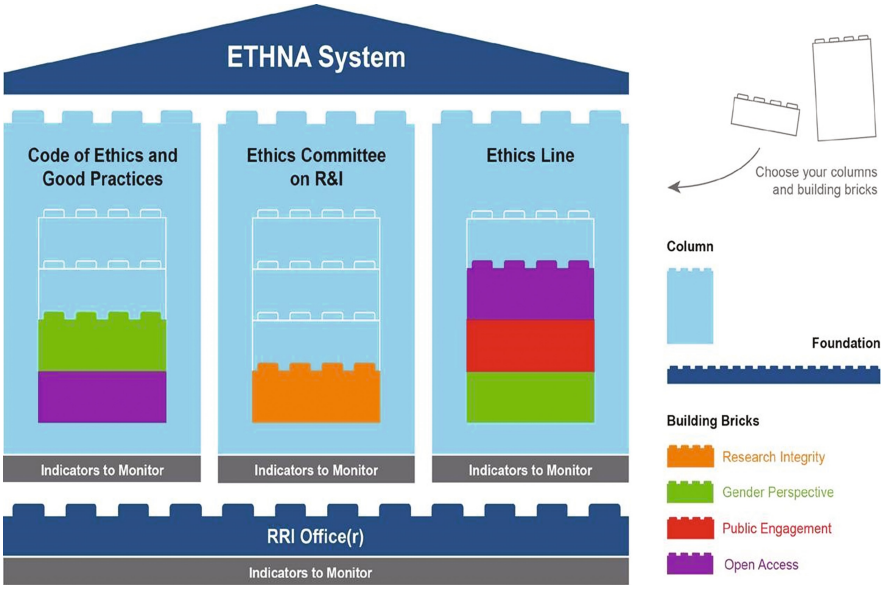


Fig. 1. Blocks for constructing the ETHNA System.

institutional factors and the adaptability provided by them in order to build the ETHNA System (see [60]).

4.3 Levels of Commitment to the ETHNA System and Support Tools to Achieve Them

Based on the analysis of the *leadership* and *base,* the organisation can also define its level of institutional commitment, depending on the capabilities and willingness of its leadership. The ETHNA System shows three levels of commitment, each one always incorporating a monitoring system with progress and performance indicators, which can be found in the support tools.

Level 1. The organisation appoints an RRI Office(r) and supports its activity. Moreover, the progress and performance indicators that monitor the institutionalisation of this building block are incorporated into the management governance.
Level 2. The organisation implements the RRI Office(r) or foundation block and some of the column blocks, incorporating one or more RRI keys or issues (research integrity, gender perspective, open access and or public engagement).
Level 3. The organisation designs and implements the RRI Office(r) and implements the three columns. The organisation applied a proactive attitude in all the RRI key areas: research integrity, gender perspective, open access, public engagement or any other area or issue that has been identified as a priority for the organisation after participation and deliberation with its stakeholders.

The ETHNA System offers, in open access, a set of support tools that make it easier for any organisation to make the appropriate decisions for the configuration of its personalised ethics governance system. These ready-to-use guides show, step by step, the decisions that the organisation should adopt so that each RPO or RFO builds its ethics governance system paying attention to its base and its leadership. The organisation with the support tools provides an example that can serve as an entry point for the institution to work through co-creation following the proposed ETHNA Lab methodology and designing its own. These examples are illustrated with good practices and examples that are the result of the experiences provided by experts and organisations that address ethics governance and RRI, either in their different dimensions or in the themes selected by the ETHNA System. It is therefore a starting point to be able to begin the path to reach the proposed level.

In short, ETHNA System offers any organisation a step-by-step *Guide for the Implementation of the ETHNA System*, as well as support tools [51].

4.4 The Building Blocks in a Monitored Continuous Improvement System

The fundamental objective of the basic structures and functions of these elements making up the organisation's ethics governance of research, as understood in the ETHNA System, is given below. For space reasons, it is not possible to go into detail about each of these elements, nor was it possible to go into depth in Sect. 3.3. on the aspects that could be covered and that would provide the themes to be worked on with content. The following is a summary of some of the key points detailed in the public reports that can be found in the results of the ETHNA System project, which are openly accessible [51].

RRI Office (R)

The ETHNA Office is called the RRI Office or RRI Officer. This is because it may take the form of an administrative structure with an ethics officer as a leader who coordinates the tools (column blocks) and promotes the alignment of the existing resources by means of the ethics governance of research and innovation aligned with RRI. Although the existence of this role would be advisable, if it does not exist there is a need to establish who assumes these responsibilities – a possible option would be the ethics committee.

Having a formal Ethical Office of Research and Innovation with a high-level executive position leading an ethical infrastructure in an organisation would be very helpful in guiding professional and institutional moral performances [61]. The creation and maintenance of such structures will show a deep commitment to the promotion of ethical and responsible behaviour in research and innovation.

Code of Ethics and Good Practices in R&I
This is a self-regulatory document that explicitly outlines the principles, values, and good practices that should guide the activity of the people involved in R&I processes, as well as the organisation's policies and programmes.

Ethics Committee on R&I
This is an internal consultation and arbitration body that acts as a forum for participation, reflection, and dialogue between the organisation's different stakeholders on R&I matters.

Ethics Line
This is a communication channel that allows all stakeholders to easily and safely send the organisation suggestions, warnings, complaints, and reports.

ETHNA System has been designed as a continuous evolution system towards RRI. For this reason, it is recommended that the results measured by the indicators (progress and performance) and the adequacy of the indicators themselves have an annual or biannual review integrated into the RRI Action Plan. The RRI Office(r) will be responsible for reviewing these indicators and making progress visible, both internally and externally. This will allow the improvement of the construction, reinforcing both the *base* and the organisation's *leadership* [51].

5 Summary and Final Remarks

ETHNA is a flexible ethics governance system designed to be implemented in RPOs and RFOs, in different contexts, i.e., universities, organisations funding research and innovation, research and development centres or innovation ecosystems. It follows two compass criteria: the ethical criteria of the "all-affected principle" and the effective criteria of "de facto governance". In this way, it responds to the four procedural dimensions of anticipation, inclusion, reflection and responsiveness of the RRI framework. It also provides a starting point on how four of the multiple agendas and key research topics can be specified in the ETHNA System, namely: research integrity, the gender perspective, open access and public engagement.

The ETHNA System offers ethics governance structures based on a system of flexible blocks that can be adapted to the needs and particular features of each organisation and their available resources. The ETHNA System allows organisations to build their own ethics governance structure for knowledge-generation and innovation processes, and make progress by continuously improving them over time.

As has already been indicated, one of the aspects to continue working on within this sustained improvement model is to address other issues with this same system, such as environmental sustainability or social justice in R&I, among other agendas or issues shown to be demands or requirements of society. In short, where research and innovation should be open to continue addressing these and other issues, together with society. In our view, the ETHNA System allows for this continuous openness to the values and social and ethical expectations of its stakeholders through its ethics governance system.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission.

References

1. Horn, L.: Promoting responsible research conduct: a south african perspective. J. Acad. Ethics **15**(1), 59–72 (2016). https://doi.org/10.1007/s10805-016-9272-8
2. Hoecht, A.: Whose ethics, whose accountability? a debate about university research ethics committees. Ethics Educ. **6**(3), 253–266 (2011). https://doi.org/10.1080/17449642.2011.632719
3. Moan, M.H., Ursin, L., de Grandis, G.: Institutional governance of responsible research and innovation. In: Ethics and Responsible Research and Innovation in Practice, LNCS, **13875**. Springer, Cham (2023)
4. Cortina, A., García Marzá, D., Conill, J. (eds.): Ashgate, Aldershot (2008)
5. Habermas, J.: Justification and Application: Remarks on Discourse Ethics. Polity Press, Boston (1993)
6. Häberlein, L., Hövel, P.: Importance and necessity of stakeholder engagement. In: Ethics and Responsible Research and Innovation in Practice, LNCS, **13875**. Springer, Cham (2023)
7. Hajdinjak, M.: Implementing RRI in a non-governmental research institute. In: Ethics and Responsible Research and Innovation in Practice, LNCS, **13875**. Springer, Cham (2023)
8. Camarinha-Matos, L.M., Ferrada, F., Oliveira, A.I.: Implementing RRI in a research and innovation ecosystem. In: Ethics and Responsible Research and Innovation in Practice, LNCS, **13875**. Springer, Cham (2023)
9. Bernal-Sánchez, L., Feenstra, R.A.: Developing RRI and research ethics at university. In: Ethics and Responsible Research and Innovation in Practice, LNCS, **13875**. Springer, Cham (2023)
10. Landeweerd, L., Townend, D., Mesman, J., Van Hoyweghen, I.: Reflections on different governance styles in regulating science: a contribution to 'responsible research and innovation.' Life Sci., Soc. Policy **11**(1), 1–22 (2015). https://doi.org/10.1186/s40504-015-0026-y
11. Stilgoe, J., Owen, R., Macnaghten, P.: Developing a framework for responsible innovation. Res. Policy **42**(9), 1.568–1.580 (2013) https://doi.org/10.1080/17449642.2011.632719
12. Stilgoe, J., Owen,R., Macnaghten, P., Stilgoe, J., Macnaghten P., Gorman M., et al.: Developing a framework for responsible innovation. Res. Policy **42**(9), 1,568–80. https://doi.org/10.1016/j.respol.2013.05.008
13. Jonas, H.: The Imperative of Responsibility. University of Chicago Press, Chicago (1984)
14. Von Schomberg, R.: A vision of responsible research and innovation. In: Owen, R., Bessant, J., Heintz, M. (eds): Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society, pp. 51–74. John Wiley & Sons Inc., Chichester, West Sussex (2013). https://doi.org/10.1002/9781118551424.ch3
15. Owen, R., von Schomberg, R., Macnaghten, P.: An unfinished journey? reflections on a decade of responsible research and innovation. J Responsible Innov. **8**(2), 217–233 (2021). https://doi.org/10.1080/23299460.2021.1948789
16. Von Schomberg, R., González-Esteban, E., Sanahuja-Sanahuja, R.: Ethical challenges and limits of RRI for improving the governance of research and innovation processes. Recer Rev Pensam i Anàlisi. **27**(2), 1–6 (2022). https://doi.org/10.6035/recerca.6750

17. Owen, R., Pansera, M.: Responsible innovation and responsible research and innovation. In: Simon D., Kuhlmann S., Stamm J., Canzler W. (eds): Handbook on Science and Public Policy, pp. 26–48. Edward Elgar Publishing, Cheltenham, UK; Northampton, MA (2019), https://doi.org/10.4337/9781784715946.00010
18. Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., Guston, D. A framework for responsible innovation. In: Owen R., Bessant J., Heintz M., (eds.): Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Chichester, West Sussex: John Wiley & Sons Inc., pp. 27–50 (2013). https://doi.org/10.1002/9781118551424.ch2
19. Rip, A.: The past and future of RRI. Life Sci. Soc. Policy **10**(1), 1–15 (2014). https://doi.org/10.1186/s40504-014-0017-4
20. Stahl, B.C., Eden, G., Jirotka, M., Coeckelbergh, M.: From computer ethics to responsible research and innovation in ICT. The transition of reference discourses informing ethics-related research in information systems. Inf. Manag. **51**, 810–818 (2014). https://doi.org/10.1016/j.im.2014.01.001
21. Beck, U., Ritter, M.: Risk Society: Towards a New Modernity. Sage Publications, London (1992)
22. González-Esteban, E.: Ética y gobernanza: un cosmopolitismo para el siglo XXI. Granada, Comares (2013)
23. Irwin A. The Politics of Talk: Coming to Terms with the "New" Scientific Governance. **36**, Social Studies of Science, pp. 299–320 (2006). https://doi.org/10.1177/0306312706053350
24. Schuijff, M., Dijkstra, A.M.: Practices of responsible research and innovation: a review. Sci. Eng. Ethics **26**(2), 533–574 (2019). https://doi.org/10.1007/s11948-019-00167-3
25. Godecharle, S., Nemery, B., Dierickx, K.: Heterogeneity in european research integrity guidance: relying on values or norms. J. Empir. Res. Hum. Res. Ethics **9**(3), 79–90 (2014)
26. González-Esteban, E.: ¿Qué tipo de reflexividad se necesita para fomentar la anticipación en la RRI? una visión ético-crítica. In: Rodriguez Zabaleta, H., Ureña López, S., Eizagirre Eizagirre, A., Imaz Alias, O. (eds.) Anticipación e innovación responsable: la construcción de futuros alternativos para la ciencia y la tecnología, pp. 229–249. Minerva Ediciones, Barcelona (2019)
27. Habermas, J., The Idea of the University: Learning Processes. New Ger Crit. 1987; (41, Special Issue on the Critiques of the Enlightenment), pp. 3–22
28. García-Marzá, D.: From ethical codes to ethical auditing: an ethical infrastructure for social responsibility communication. El Prof la Inf. **26**(2), 268–276 (2017). https://doi.org/10.3145/epi.2017.mar.13
29. García-Marzá, D., Fernández Beltrán, F., Sanahuja, R.: Ética y Comunicación en la gestión de la investigación e innovación responsable (RRI). El papel de las unidades de cultura científica y de la innovación (UCC+I). Castellón de la Plana: Publicacions de la Universitat Jaume I (2017). https://doi.org/10.6035/Humanitats.2017.52
30. Flipse, S.M., van der Sanden, M.C.A., Osseweijer, P.: The why and how of enabling the integration of social and ethical aspects in research and development. Sci. Eng. Ethics **19**(3), 703–725 (2013). https://doi.org/10.1007/s11948-012-9423-2
31. Özdemir, V.: Towards an ethics-of-ethics for responsible innovation. In: von Schomberg R., Hankins J., (eds): International Handbook on Responsible Innovation: A Global Resource. Cheltenham, UK; Northampton, MA: Edward Elgar Publishing, pp. 70–82 (2019). https://doi.org/10.4337/9781784718862.00011
32. European Commission. Responsible Research and Innovation. Europe's ability to respond to societal challenges, p. 4 (2014). https://op.europa.eu/en/publication-detail/-/publication/2be36f74-b490-409e-bb60-12fd438100fe
33. Guston, D.H.: Understanding "anticipatory governance." Soc Stud Sci. **44**(2), 218–242 (2014). https://doi.org/10.1177/0306312713508669

34. Guston, D.H., Sarewitz, D.: Real-time technology assessment. Technol Soc. **24**(1–2), 93–109 (2002). https://doi.org/10.1016/S0160-791X(01)00047-1
35. Eizagirre, A.: Investigación e innovación responsables: retos teóricos y políticos. Sociol Probl e Práticas. **2017**(83), 99–116 (2016). https://journals.openedition.org/spp/2713?lang=pt
36. Braun, R., Griessler, E.: More democratic research and innovation. J. Sci. Commun. **17**(3), 1–7 (2018). https://jcom.sissa.it/article/pubid/JCOM_1703_2018_C04/
37. Burget, M., Bardone, E., Pedaste, M.: Definitions and conceptual dimensions of responsible research and innovation: a literature review. Sci. Eng. Ethics **23**(1), 1–19 (2016). https://doi.org/10.1007/s11948-016-9782-1
38. Forsberg, E.-M., Shelley-Egan, C., Ladikas, M., Owen, R.: Implementing responsible research and innovation in research funding and research conducting organisations—what have we learned so far? In: Governance and Sustainability of Responsible Research and Innovation Processes. SRIG, pp. 3–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73105-6_1
39. van Hove, L., Wickson, F.: Responsible research is not good science: divergences inhibiting the enactment of RRI in nanosafety. NanoEthics **11**(3), 213–228 (2017). https://doi.org/10.1007/s11569-017-0306-5
40. Owen, R., Macnaghten, P., Stilgoe, J.: Responsible research and innovation: from science in society to science for society, with society. Sci. Public Policy **39**(6), 751–760 (2012). https://doi.org/10.1093/scipol/scs093
41. Von Schomberg, R.: Towards Responsible Research and Innovation and Communication Technologies and Security Technologies Fields. Publications Office of the European Union. Luxembourg (2011). https://op.europa.eu/en/publication-detail/-/publication/60153e8a-0fe9-4911-a7f4-1b530967ef10
42. Von Schomberg, R.: Prospects for technology assessment in a framework of responsible research and innovation. In: Dusseldorp M., Beecroft R., editors. Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methode. Springer, Wiesbaden, pp. 39–61 (2012). https://doi.org/10.1007/978-3-531-93468-6_2
43. García-Marzá, D., Fernández Beltrán, F., Sanahuja, R., Andrés, A.: El diálogo entre ciencia y sociedad en España. Experiencias y propuestas para avanzar hacia la investigación y la innovación responsables desde la comunicación. Castellón de la Plana: Publicacions de la Universitat Jaume I (2017). https://doi.org/10.6035/FCT.16.11124.2018
44. Stilgoe, J., Lock, S.J., Wilsdon, J.: Why should we promote public engagement with science? Public Underst Sci. **23**(1), 4–15 (2014). https://doi.org/10.1177/0963662513518154
45. Von Schomberg, R.: From the ethics of technology towards an ethics of knowledge policy & knowledge assessment. Publications Office of the European Union. Luxembourg (2007). https://op.europa.eu/en/publication-detail/-/publication/aa44eb61-5be2-43d6-b528-07688fb5bd5a
46. Honneth, A.: The Critique of Power: Reflective Stages in a Critical Social Theory. MIT Press, Cambridge, MA (1991)
47. Gianni, R.: Choosing freedom: ethical governance for responsible research and innovation. In: von Schomberg, R., Hankins, J., (eds): International Handbook on Responsible Innovation: A Global Resource. Cheltenham, UK; Northampton, MA: Edward Elgar Publishing, pp. 49–69 (2019). https://doi.org/10.4337/9781784718862.00010
48. Wynne, B.: Public uptake of science: a case for institutional reflexivity. Public Underst Sci. **2**, 321–337 (1993). https://doi.org/10.1088/0963-6625/2/4/003
49. Taebi, B., Correljé, A., Cuppen, E., Dignum, M., Pesch, U.: Responsible innovation as an endorsement of public values: the need for interdisciplinary research. J. Responsible Innov. **1**(1), 118–124 (2014). https://doi.org/10.1080/23299460.2014.882072
50. ETHNA System Project. https://ethnasystem.eu/about-ethna/the-project/

51. González-Esteban, E., et al.: The ETHNA system – a guide to the ethical governance of RRI in innovation and research in research performing organisations and research funding organisations. Deliverable 4.2 Draft concept of the ETHNA System. ETHNA Project [872360] Horizon (2020). Zenodo. https://doi.org/10.5281/zenodo.6532789
52. European Science Foundation, ALLEA – All European Academies. The European Code of Conduct for Research Integrity, pp. 1–20 (2011). https://allea.org/code-of-conduct/
53. Habermas, J.: The Theory of Communicative Action, vol. 2. Beacon Press, Boston (1987)
54. Stahl, B.C.: Morality, ethics, and reflection: a categorization of normative is research. J. Assoc. Inf. Syst. **13**(8), 636–656 (2012). https://doi.org/10.17705/1jais.00304
55. Rehg, W.: Discourse ethics for computer ethics: a heuristic for engaged dialogical reflection. Ethics Inf. Technol. **17**(1), 27–39 (2014). https://doi.org/10.1007/s10676-014-9359-0
56. Sorensen, E.: The metagovernance of public innovation in governance networks. In: Paper presented at the Policy and Politics Conference in Bristol (2014)
57. Sorensen, E., Torfing, J.: Making governance networks effective and democratic through metagovernance. Public administration. Public Adm. **87**(2), 234–58 (2009). https://www.bristol.ac.uk/media-library/sites/sps/migrated/documents/sorensonthemetagovernanceofpublicinnovation.pdf
58. Kooiman, J.: Social-political governance: overview, reflections and design. Public Management and international journal of research and theory. Public Manag Int. J. Res. Theory. **1**(1), 67–92 (1999). https://doi.org/10.1080/14719037800000005
59. Hovdal Moan, M., Ursin, L., González-Esteban, E., Sanahuja Sanahuja, R, Feenstra, R., Calvo, P., et al.: ETHNA system: literature review and state of the art description. Mapping examples of good governance of research and innovation (R&I) related to responsible research and innovation (RRI), in Higher Education, Funding and Research Organisations (HEFRCs) (2022). https://ethnasystem.eu/wp-content/uploads/2021/02/ETHNA_Report_state-of-the-art.pdf
60. Szüdi, G., Lampert, D., Hajdinjak, M., Asenova, D., Alves, E., Bidstru, M.V.: Evaluation of RRI institutionalisation endeavours: specificities, drivers, barriers and good practices based on a multi-stakeholder consultation and living lab experiences. In: Ethics and Responsible Research and Innovation in Practice, Springer, Cham (2023)
61. Adobor, H.: Corporate ethics officers. In: Poff, D.C., Michalos, A.C., (eds): Encyclopedia of Business and Professional Ethics. Springer Nature, Switzerland (2019). https://doi.org/10.1007/978-3-319-23514-1_66-1

Importance and Necessity of Stakeholder Engagement

Lisa Häberlein(✉) and Philipp Hövel

EUREC (European Network of Research Ethics Committees), Bonner Talweg 57, 53113 Bonn, Germany
{haeberlein,hovel}@eurecnet.eu

Abstract. This chapter addresses the importance and necessity of stakeholder engagement for ethics governance. It starts with a more general look at stakeholder engagement and describes how a comprehensive definition of the otherwise commonly used catch-all term of stakeholders is important for successful stakeholder engagement. This is followed by a description of the ETHNA System's stakeholder engagement strategy. A step-by-step approach serves as a guide for stakeholder engagement in responsible research and innovation (RRI) activities. This chapter offers guidance on how to find answers to the questions of who relevant stakeholders are, why they should be involved, what their contribution can be and how they can be involved in the most effective way. It also explains how to identify stakeholders' values, needs and expectations using different tools. Finally, lessons learned from implementing the stakeholder engagement strategy are discussed to make the ETHNA System more useful to others. Specific guidelines on stakeholder engagement are used as a reference in this text.

Keywords: Stakeholder Engagement · Stakeholder Mapping · Ethical Governance · RRI · R&I · Guidelines · Lessons Learned

1 Towards Input and Output Legitimacy

Engaging a broad range of stakeholders is widely recognized as crucial to responsible research and innovation (RRI) activities and good research governance. However, existing concepts of stakeholder engagement offer only limited guidance on who relevant stakeholders are, why they should be engaged, what they can contribute, and how they can best be involved [1]. In this chapter, we bring together the scattered structures and outline how higher education, funding, and research centers (HEFRCs) and other organizations can find answers to the questions of the "who", "why", "what" and "how" [1–3].

Given that the ETHNA System can only be successful in influencing the innovation process and ensuring that its outcomes are legitimate, effective, and efficient, and that the resulting innovations avoid harm and do good for society and the planet [1] input and output legitimacy must be ensured. Therefore, this chapter discusses how the ETHNA

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 38–53, 2023.
https://doi.org/10.1007/978-3-031-33177-0_3

System facilitates an inclusive and transparent participatory RRI process through stakeholder engagement, reflecting and embracing the goals as well as the means and social acceptability of innovations [1–3].

As an internal management and procedural system for RRI, the ETHNA System introduces a new formal organizational structure. Considering that RRI has its origins in the risk assessment of scientific innovations, the European Commission (EC) describes RRI as "[…] an ongoing process of aligning research and innovation with society's values, needs and expectations" [4]. In this context, socially useful, meaningful, or desired outcomes and effects should be the guiding principles of research and research funding. To enable societal actors and innovators to respond to each other in terms of the (ethical) acceptability, sustainability and societal desirability of the innovation process and its outcomes, RRI processes need to be as transparent and interactive as possible [1, 5]. The responsibility associated with research and innovation (R&I) is thus expressed in involving society and responding to its needs. Where fundamental ethical dimensions and questions of social acceptance are left out or considered too late in the research process, irresponsible innovation can be the result. Thus, involving relevant stakeholders in deliberative participation processes at an early stage guarantees that societal values, needs, and expectations are effectively identified, discussed, and considered.

Ethical governance towards RRI is implemented in the ETHNA System through multi-stakeholder governance in participatory processes (to learn more about the ETHNA System in general, see [6]). Participatory processes can be understood as an effective means to gather perspectives on complex issues, to share relevant research findings and/or to reach an agreement or build consensus in cases where polarized views and conflicts of interest exist [7]. This can build trust between science and society and promote social acceptance of R&I, especially publicly funded research [8]. Input and output legitimacy thus revolves around the actors who aim to "make the innovation process socially desirable" and contribute to the operationalization of the normative ideals of inclusion and social legitimacy in R&I [1, 5].

2 Dealing with a Catch-all Term

Stakeholder engagement research that draws on stakeholder theory typically focuses on the conceptual and theoretical development and organizational and societal benefits of stakeholder engagement. Kujala et al. argue that stakeholder engagement is a widely used but nevertheless unspecific concept, as far as a mutual understanding of the essential elements of stakeholder engagement is lacking [9]. Yet the importance of stakeholder engagement in activities such as those related to R&I, strategic planning and decision-making, knowledge creation, or progress monitoring to report on RRI, all of which play a significant role in the implementation of an ETHNA System, is undisputed. Inclusiveness, participation, and engagement of a wide range of stakeholders are also key to RRI according to empirical studies [1, 10–13].

However, diverging definitions of stakeholder engagement make the endeavor challenging. While the currently most prominent definition describes stakeholder engagement primarily as a morally neutral practice that an organization undertakes to involve stakeholders in organizational activities [9, 14], there are other definitions that either

elaborate a stronger focus on the moral, strategic, or pragmatic components of stakeholder engagement. The ETHNA System places the moral component at the center of stakeholder engagement by addressing ethical principles to suggest stakeholder-oriented responses to strategic ethical challenges in organizational governance (see [6]).

These may include issues related to expectations of legitimacy, trust and fairness in stakeholder relationships and the organizational structures, as well as morally desirable outcomes regarding institutional change, such as improved space and communication channels for expression, participation and dialogue or better orientation and compliance with ethics regulations.

Braun and Starkbaum criticize in this context that stakeholders are used in most RRI discourses as a catch-all term for societal actors without making clear who they are, why their participation is important, and what exactly they might contribute to the R&I process. Thus, in practice, it is often not clear why one and not the other stakeholder is invited and for what purpose [1, 15]. Aspects which are specified in the ETHNA System by means of concrete guidance for organizations, considering their institutional needs [16–18].

Thus, in the development of an ethics governance system, the nuanced understanding of the moral component means a discussion about the organizational needs under the premise of philosophical foundations for stakeholder engagement, such as discourse ethical principles. Discourse ethics is made strong in the ETHNA System by highlighting criteria for reflection, participation, and dialogue between various stakeholders. This approach is a major step towards understanding stakeholders in their full complexity. Twelve stakeholder inclusion principles, considering this philosophical tradition, are incorporated into the ETHNA System to support organizations in the implementation process.

Although normative stakeholder theory plays a significant role in the ETHNA System, the pragmatic and strategic components of stakeholder engagement are also relevant. Descriptive stakeholder theory assumes that a variety of different stakeholders represent different values, needs and expectations, arguing that these different viewpoints on potential societal impacts of R&I should be incorporated into deliberative RRI activities. This is not least with the aim of building and sustaining relationships for organizational and societal change through context- and time-related collaborative activities. Instrumental stakeholder theory, on the other hand, places a strong focus on strategic organizational challenges to better achieve organizational goals. This approach receives great attention in relation to management and strategy, making purpose- and goal-oriented aspects particularly strong. Thus, strategic activities are based on improving benefits and reducing risks, to foster understanding of the importance of implementing an ETHNA System, ensuring commitment and higher levels of achievement.

Since, the moral, pragmatic, and strategic components, all play a role in the implementation of an ETHNA System, a comprehensive definition of stakeholder engagement that refers to the goals, activities, and effects of stakeholder relationships in a moral, strategic, and pragmatic way, deserves special appreciation [9].

3 Guiding the Implementation in Practice

To put stakeholder engagement on solid theoretical and methodological grounds, the ETHNA System provides three guidelines that help HEFRCs and other organizations in the engagement process [16–18].

Before starting to engage stakeholders, their topology needs to be identified and analyzed. To ensure a diversified group of stakeholders, Schütz et al. recommend the quadruple helix model (QHM), which is also followed in the ETHNA System. By grouping stakeholders in different sectors, greater diversity is ensured. Schütz's four identified sectors are: science, policy, industry, and society [19] and are adapted for the ETHNA System project to the objectives and needs of HEFRCs as follows:

- Research, innovation, funder community,
- Policy makers,
- Business and industry, and
- Civil society.

Thereby, the stakeholder groups are not fixed but rather fluent and depending on the context. The same stakeholder might be a business representative in one case and a member of civil society in another. When identifying and mapping stakeholders, it is also important to not just analyze them by sector but also by interest and by location [20]. To systematically engage relevant stakeholders for RRI activities, the ETHNA System recommends the following six steps:

1. identify,
2. analyze,
3. map,
4. prioritize
5. select, and
6. recruit relevant actors [16].

The first step in systematic stakeholder engagement is to identify relevant stakeholders. Starting with a brainstorming session within the organization or institution, different techniques can be used to create a diverse list of actors. Relevant stakeholders can be identified by reviewing address books, engaging one's own or the institution's network, or by reviewing relevant literature. Conferences, workshops, forums etc. might also reveal stakeholders that are or have already been working on related topics or show interest in engaging with RRI issues. Once initial stakeholders have been identified, they can also be asked to recommend further potentially interested people or use their own networks and contacts. In collaborating with an ever-expanding group, the stakeholder list might grow fast. However, when adding stakeholders to the list, some guiding questions should be considered:

- Who might be affected by the RRI activity?
- Who deals with the respective issue of the RRI activity and who qualifies because of position or influence?

- Who are hidden, invisible or indirect stakeholders?

The compiled stakeholder list can be considered as a living document that can change over time. Of course, requirements can vary during the implementation process, so the list can grow by adding stakeholders with a new perspective, or by removing stakeholders from the list that no longer fit the requirements. Nevertheless, the stakeholders already identified must be analyzed to examine their role in the RRI activity. Therefore, the ETHNA System project created a template to analyze, e.g., the willingness or influence of certain stakeholders (see Table 1).

Table 1. Stakeholder analysis example

	Stakeholder	Contribution	Legitimacy	Willingness	Influence	Necessity
1	Research, Innovation, Funder Community	High	High	High	High	High
2	Business and Industry	High	High	Medium-Low	High	High-Medium
3	Policy Makers	High	High	High	High	High
4	Civil Society	Low	High	Medium-Low	Low	Medium

After analyzing stakeholders, they can additionally be mapped, e.g., to visualize their expertise in comparison to their willingness to participate in the RRI activity. This ensures that an adequate overall representation of the stakeholders is created and that they can participate in activities that best suit their knowledge, skills, attitudes, and values, as well as their needs and expectations and at the same time correspond to the interests of the RRI activity. The overarching goal is, on the one hand, to benefit from the diverse knowledge and expertise of stakeholders and, on the other, to keep stakeholders motivated and engaged and let them see the outcome and impact from their contribution. Guiding questions to consider while mapping stakeholders, are e.g.:

- Are the stakeholders well networked and well known in their respective fields?
- Are there personal or business-related connections that can be used or revived?
- Are the actors familiar with relevant topics, such as research integrity (RI), gender perspective, open access (OA), or public engagement?
- Will they contribute to the implementation of an ETHNA System in a positive way?

According to Moan et al. “an ethics of involvement concerns, not the question of who should be involved in RRI processes and why, but the question of how the persons involved should be involved” [21]. In this sense, stakeholders should be categorized in terms of collaboration, involvement, consultation, and information [22]. This step allows

prioritizing stakeholders according to the level of their potential engagement (as shown in Table 2).

Table 2. Stakeholder participation template [16]

Higher Priority Stakeholders (Individuals, Companies, Organizations, Communities, Associations, Parties etc.)	**Level of Potential Engagement**			
Research, Innovation, Funder Community	**Collaboration**	**Involvement**	**Consultation**	**Information**
Business and Industry				
Policy Makers				
Civil Society				

This refers to a bigger theoretical discussion which was elaborated in more detail in the PE2020 project [23]. In short, a distinction is made between one-dimensional forms of public engagement, namely public communication, public consultation and public activism, and two-dimensional forms, public participation, and public deliberation. In one-dimensional communication, stakeholders are mainly educated and informed, while there is no mechanism for feedback. In a two-dimensional dialogue "the exchange is accompanied by a debate in which knowledge can be acquired and applied at the same time" [23]. With the latter approach, reflection and responsiveness are ensured and different perspectives are considered. Therefore, in areas where stakeholders are directly affected by the research, deliberative methods are recommended. This is firstly, to get the most out of the process so that the RRI activity can be successful, and secondly, to ensure that all stakeholders have a say and benefit from the process as much as the project or organization does. According to Rask et al. "there has been a shift of PE [public engagement] from traditional models of public communication and consultation, where dialogue between decision makers and the public is narrow and restricted, to public deliberation where such dialogue is intensive and influential" [24].

After stakeholders have been identified, analyzed, mapped, and prioritized, the next step is stakeholder selection and recruitment. All the information gathered in the previous steps can be summarized in a stakeholder list with columns on stakeholder group, contact details, level of involvement, link to relevant topics, such as research integrity, gender perspective, open access and public engagement. As mentioned before, this stakeholder list should be a living document that can be assessed regularly [25] and evolve over the course of the RRI activity [20]. Finally, the relevant stakeholders are contacted to engage them in the activity. This should be done by being as specific as possible, e.g., about the

aims and expectations, the roles of the stakeholders, the impact their engagement can have, or the benefits of their participation.

Once stakeholders have been mapped and a stakeholder list has been compiled[1], it is necessary to involve them in the implementation process. For that purpose, the ETHNA System recommends twelve principles for deliberative participation that allow involving relevant stakeholders as effectively as possible and with consideration of fundamental ethical values[2]:

1. Intentions and expectations need to be clear right from the beginning. This means starting as early as possible to specify the RRI activity and share information with stakeholders. Also, the extent to which stakeholders might have an impact on outcomes must be clarified. In this way, the first step towards a culture of openness, transparency, and participation is promoted.
2. Sufficient resources in terms of time, skills, and funding for the engagement processes must be ensured. For this purpose, systematic planning and budgeting is important.
3. Persons and institutions must be mapped alike. Stakeholders can be either individuals or groups that might affect or be affected by the ethical governance system.
4. Diversity must be embraced so that stakeholders representing different interests and groups contribute a wide range of perspectives.
5. Directly and indirectly affected stakeholders should be part of the RRI activity.
6. Values, needs, expectations, interests, and concerns of stakeholders from the research, innovation and funding community, business and industry, politics and civil society should be covered and as many perspectives as possible should be included.
7. Local factors must be considered, and it should be reflected in whether the identified stakeholder is a national or international actor. E.g., a city, region, country, neighboring country, or international context can be assigned to them.
8. Different stakeholder groups, RRI key areas, levels of engagement or process dimensions for RRI should be spanned. This reveals answers to the questions of the "who", "why", "what", and "how".
9. Supporting stakeholders in engaging in a discourse is crucial. Collaboration, networking, broader participation, and co-operation in relation to engagement with RRI should be encouraged.
10. Engagement methods and techniques that are appropriate to the aims of the RRI activity should be employed and a variety of deliberative engagement techniques should be demonstrated to prevent stakeholder fatigue. To respond to different

[1] According to Creighton, internal and external stakeholders should be distinguished. When implementing an ethics governance system at organizational level, it can be assumed that internal stakeholders have very specific values, needs and expectations regarding their everyday activities. External stakeholders, on the other hand, may be able to contribute new perspectives or report on experiences not (directly) related to the organization. Their needs are likely to be different from those of internal stakeholders [20].

[2] The ETHNA System has adopted and adjusted these principles from other sources [20, 26–28].

participants, the topics addressed should be oriented towards their knowledge, experience, skills, and controversy.
11. The six steps of the systematic stakeholder engagement process should be evaluated.
12. Academic freedom must be protected in the RRI activity, which is valuable and even a fundamental right in some communities.

By following these twelve principles and the six steps for systematic stakeholder engagement, the complex process of stakeholder engagement can be systematized and made easier to manage. To this end, organizations are encouraged to address the questions of the "who", "why" "what" and "how" as a means of maximizing the effectiveness of engaging stakeholders. These aspects were tested and applied by implementing organizations during the ETHNA System project's lifetime, reportedly leading to success stories and challenges discussed in the next section of this chapter.

When striving to align research with society's values, needs and expectations, and following the European Commission RRI definition [4], it is important to know what society's needs are. These can be determined in diverse ways. As a survey among a variety of European networks conducted during the project showed, and many other actors – research funding organizations (RFOs), political institutions and researchers – confirmed, research performing organizations (RPOs) themselves are responsible for responding "to societal needs when promoting R&I" [16]. Six pressing societal needs to which R&I might respond, emerged from the abovementioned survey in combination with a secondary desk research:

1. Sustainability/protecting land and oceans
2. Data Protection/privacy/protection of human rights
3. New Technologies/AI/robotics
4. Health (including mental health and well-being)
5. Food/farming
6. (Drinking) Water

To keep these needs in mind, compare or even expand them by incorporating stakeholder perspectives, the ETHNA System recommends HEFRCs and other organizations to use the following three techniques and resources:

- Eurobarometer
- Funding programs and calls
- Foresight approaches.

Eurobarometer surveys are long-term series of public opinion polls on a variety of subjects [17]. Insights into current trends and developments can be gained as the expectations of civil society of the European Union´s (EU) policies are gathered (https://www.europarl.europa.eu/at-your-service/en/be-heard/eurobarometer).

Moreover, it is possible to access the results and actively participate in improving them further. Another option is to review calls and projects that have recently been funded, e.g., by Horizon Europe, as the topics they deal with are most likely to respond to or at least reflect societal needs. A third method can be foresight approaches. They

are "a systematic, participatory process for gathering knowledge about the future and creating visions" [16]. Again, it is possible to either visit the European Foresight Platform (http://foresight-platform.eu/) or to conduct one's own foresight exercise with relevant stakeholders. How this can be done is explained in detail in the ETHNA System guide "Gauging the potential societal contributions of research and innovation" [17].

The stakeholder engagement strategy in the ETHNA System project can be resumed as follows: "Involving wider populations in participatory events offers opportunities to define recommendations for R&I. By considering different options for complex issues, better informed decisions can be made" [18]. It is also important to involve various stakeholders, experts and (international) networks to gather insights on the best ways to keep on track with (changing) societal needs and how to assess and address them.

4 Practical Advice and Lessons Learned

4.1 Addressing the Pitfalls

Publishing not only positive, but also negative results is at the heart of the scientific enterprise. However, the focus is on positive results, which is unfortunate as this means that a lot of highly useful information does not reach the public. Also, regarding the issue of stakeholder engagement, there is an overwhelmingly positive attitude towards the stakeholder engagement approach, with the result that the "dark side of stakeholder engagement" is overlooked [9]. This is evident through extensive literature on the goals and expectations on the organizational and stakeholder sides. Problematic aspects of engagement, on the other hand, are addressed by only a few authors.

As it is essential for RRI to ensure that resources are used as productively as possible and to learn from those who have already undertaken a similar path, the ETHNA System project shares experiences from the implementation process. This chapter presents several RRI activities where stakeholder engagement is embedded in the implementation of the ETHNA System and discusses the challenges and difficulties encountered in this process. This provides a better understanding of the challenges and lessons learned from the implementation process of the different organizations and institutions and helps future implementers of the ETHNA System to harmonize RRI activities and to identify and keep an eye on potential pitfalls early on. This should provide opportunities to address, e.g., challenges with prominent levels of participation and stakeholder fatigue, resistance from different stakeholders, lack of time, or the difficulty in bringing external stakeholders on board.

4.2 Stakeholder Engagement Experiences from the Implementing Process

The ETHNA System, "a system of ethical governance promoting RRI in organizations, thereby encouraging them to consider the consequences of their activities and incorporate society's expectations into their work" [29] was implemented by six different organizations:

- UJI (University Jaume I), Spain, Higher Education

- NTNU (Norwegian University of Science and Technology), Norway, Higher Education
- ARC Fund (Applied Research and Communications Fund), Bulgaria, Research Centre
- Harno (Education and Youth Board of Estonia), Estonia, with competences as Research funding Organization
- UNINOVA-CTS (Instituto de Desenvolvimento de Novas Tecnologias – Center of Technology and Systems), Portugal, Innovation Ecosystem
- ESPAITEC (Parc Científic Tecnològic i Empresarial), Spain, Innovation Ecosystem

All implementing partners conducted participatory deliberative activities such as workshops, interviews or focus groups to engage with relevant stakeholders. Thereby, they were guided by three documents on stakeholder engagement: stakeholder mapping [16], potential societal contributions of research and innovation [17] and stakeholder involvement [18], developed during the ETHNA System project. These guides systematized the stakeholder engagement process and provided support on how to conduct workshops, use deliberative methods and keep track of societal needs. While implementing the stakeholder guidelines, organizations used different methods and had various experiences, outlined in more detail in the respective project outcomes [30, 31].

Considering diverse institutional needs, all HEFRCs implemented various aspects of the ETHNA System according to its building block system [29]. However, all appointed an RRI officer [30, 31], responsible for the implementation of the ETHNA System, develops the action plan and communicates it with the institution's internal stakeholders as well as with external actors. This step serves as the foundation block of the implementation [29].

In general, there are strong similarities in the reports of stakeholder engagement experiences despite the differences between the implementing organizations, e.g., in terms of type of the institution, size and corresponding capacities and resources, location and prevailing research landscape, or the institutional regulations already in place, potentially facilitating the implementation of an ETHNA System. However, it is striking that the experiences of engagement at the internal and external levels are perceived very differently.

Whereas the engagement of internal stakeholders is largely described as positive, implementers' reports on the involvement of external stakeholders are rather poor. Challenges were more apparent on the external level than on the internal, with difficulties in recruiting and actively involving external stakeholders, especially regarding smaller companies and start-ups, being mentioned by almost all implementing organizations. Specifically, ESPAITEC found the involvement of external stakeholders to be a barrier due to existing time and resource constraints [30]. Planning the participation process in terms of available time and resources can be a major task. However, if relevant stakeholders have been identified and are willing to engage, the use of deliberative methods that match their interests can turn the endeavor into a success [18]. Many deliberative techniques can be used to prevent stakeholder fatigue [28]. Harno, in turn, has conducted various RRI activities with internal and external stakeholders, such as interviews and workshops [30], and found that not being able to have the first meeting face-to-face, due to the COVID-19 pandemic, proves to be a barrier to further engagement. Introducing an ethics governance system that provides new tools, such as an ethical code, an

ethics committee, or an ethics hotline, could only be sketched out in web conferences. A face-to-face exchange, on the other hand, is seen as a more effective way to bring the ETHNA System closer to stakeholders and to clarify open questions or stimulate an open discourse. Face-to-face encounters are more likely to build, maintain or renew trusting relationships and networks than online meetings. Technology might be an obstacle to addressing difficult topics and expressing one's own perspective, or even concerns. Of course, non-verbal communication can also play a key role in this regard.

Interesting enough, feedback from implementing organizations differs from the reported experiences of internal and external stakeholders themselves. Internal stakeholders involved in RRI activities gave feedback that the process inspired them to develop further RRI key plans and make a stronger commitment to ethical governance. One stakeholder stated that this experience would lead to better implementation in the future [30]. External stakeholders' impressions of their involvement were also positive. They considered both the ETHNA System and their involvement in the implementation process as an insightful experience from which they could learn [30]. Stakeholders, whether internal or external, once involved in the engagement process, perceive their engagement in a positive way. Thus, the challenges that implementing organizations face in stakeholder engagement differ from the perceptions of stakeholders who are involved in the process or might refer to activities that take place in the systematic six-step engagement phase (as described in the previous section), in which stakeholders are not themselves involved. Thus, stakeholder engagement cannot be assessed as a high-barrier process overall.

From the perspective of implementing organizations, the involvement of external stakeholders, e.g., in UNINOVA, is seen as a significant contribution to a more effective implementation process, although it is important to recognize that their involvement is not self-evident but must be considered a process that needs to be carefully planned and conducted. This includes highlighting the benefits of participation and what stakeholders can gain from their willingness to participate. ARC Fund notes in this regard, that internal stakeholders were very busy and faced time constraints but at the same time were very interested in sharing their ideas, suggestions, concerns, and perspectives on potential impacts and issues that might arise in the implementation of an ETHNA System - especially if there is no established process or forum yet for expressing their thoughts. ARC Fund concludes that engagement is a two-way process, where something must be offered to receive something.

To dive deeper into the stakeholder perspective and learn from their feedback on the engagement process they were involved in workshops conducted by the Danish Board of Technology (DBT) [30, 31].

4.3 Reluctance from the Top Management

This chapter already discussed the role of strategic stakeholder engagement in the ETHNA System project. In that regard, implementing organizations mentioned, e.g., reluctance from the top management as a major problem. UJI reports that the involvement of the top management is crucial for compliance and sustainable implementation. Therefore, everyone in the organization must be on board and willing to commit to the ETHNA System [30]. Top management support would mean, among other things, that the organization's management understands the importance of implementing an ETHNA

System and is personally committed to it. This would promote an environment in which the ETHNA System, even in existing hierarchical structures, is recognized and appreciated, so that higher levels of achievement can be realized. The support of the top management is thus a factor for strategic implementation that would help to drive the direction and sustain the project by creating the appropriate RRI culture which promotes openness, transparency, and participation. This support is a factor that implementing organizations cannot easily control themselves. However, the fact that the implementation process of the ETHNA System includes deliberative processes with relevant stakeholders indicates that also interactions with the top management should adopt this format. It is not enough to reject a specific argument only because of a powerful position in management. According to the discourse ethical tradition, arguments are requested, expressed with respect, heard by others, and further discussed.

4.4 Reluctance from Early-Career Researchers

Yet, it is not only top-down reluctance that can hinder implementation and pose a major challenge. Reluctance of early-career researchers can also leave implementers puzzled. While it may be assumed that it is senior researchers who want to stick to established structures, the experience of the implementing organizations shows that it is earlier-career researchers who are not willing to implement an ethics governance system. Have experienced researchers already identified gaps in existing structures or begun to see the need for improvement? Either way, it is important to have stakeholders from all career levels on board. Early-career researchers represent the future of the scientific community and are the ones who will pioneer the next set of research innovations. In fact, early career researchers are the ones who will be the next generation of leaders in senior positions and form the top management of the organization. Among them are new talents with valuable knowledge, skills, attitudes, and values for implementing and sustaining an ETHNA System, so their concerns should be considered in a respectful forum, and it should be strived for finding common ground.

4.5 Conflicting Interests

The fact that some ethical recommendations of the ETHNA System might not be in line with already existing organizational procedures is a particularly delicate issue and precedes the two previous barriers. This is because when there are conflicts of interest, they get in the way of the implementation process and may also prevent the engagement of stakeholders from whatever career stage. Establishing a set of basic principles and standards to be used as a practical framework for managing and resolving conflicts of interest in accordance with best practices might be valuable. Such elements might, e.g., be included in the Code of Ethics and Good Practice, to help identify and manage issues that (repeatedly) lead to conflicts.

Thanks to the flexible structure of the ETHNA System, institutions can pick certain aspects from the toolbox and start building their own ethics governance system [29]. As in the case of NTNU, it made more sense, especially regarding efficiency and effectiveness, to build on, adapt and reassemble already existing structures [30, 31]. It is important to remember that changes should be approached cautiously. Asking for too much at once

may scare off relevant stakeholders by overwhelming them. If everyone speaks their mind during the implementation and articulates their needs and expectations, the results might be mutually beneficial and reveal insights that were not expected.

4.6 Reluctance to Adopt Ethical Principles

An important fact that implementers have learned from the engagement process is that the assumption that stakeholders are willing to adopt ethical principles is a false one [30]. The reason for existing reluctance could be that the idea of implementing an ethics governance system is accompanied by concerns about more work, more responsibility, more paperwork and so on. Even if it is a less pleasant insight, that stakeholders are rather reluctant to establish and adhere to ethical principles, it is important that they can express their concerns and worries. Therefore, it is crucial to establish an atmosphere of trust between the institution and stakeholders. In that way, they can be encouraged to exchange ideas and debates on various viewpoints and express their concerns. The ideal typical criteria for stakeholder dialogue developed during the project "can help to ensure that all stakeholders who might be affected by the RRI activity […] can accept the R&I process in a rational discourse" [16]. Besides the inclusion of all relevant stakeholders and the creation of a comfortable atmosphere, it is important to empower communication by involving "hidden", "indirect", and "invisible" stakeholders and keep in mind that minority groups are often ignored in participative approaches like these. It must be ensured that all stakeholders are heard right from the beginning and that communication is open and transparent. Therefore, a non-hierarchical dialogue can help to assess and address the reasons for hesitance in adapting ethical principles.

4.7 Lack of Awareness of Existing Regulations

Another hurdle, which follows on from the one mentioned above, is the lack of awareness and dissemination of ethical codes and other relevant regulations on an institutional level. A survey of various international networks shows that a code of ethics is key to greater accountability for R&I [17]. Accordingly, such documents provide a suitable framework for self-regulation in scientific and academic disciplines and for research environments facing new challenges. They are therefore a useful tool to support researchers and research institutions in conducting research at the highest level and help prevent misconduct by promoting best practices in R&I and providing practical guidance to the research community [17]. When developing a Code of Ethics and Good Practice, it is advised to involve stakeholders, e.g., in deliberative workshops or reviews of drafts, to include their views and make the documents useful for them.

5 Ways Ahead

By naming and acknowledging not only success stories but also challenges in participatory stakeholder activities, the first step has been taken to support those interested in implementing an ETHNA System and addressing RRI issues through systematic and

well-planned stakeholder engagement [17, 18]. Taking seriously the difficulties experienced by any kind of stakeholder, be they internal or external, further steps can be taken to strengthen the awareness of importance and promotion of the implementation of responsible governance. In particular, the following actions are recommended:

- Training early-career researchers on RRI issues. In this regard the guidelines on potential societal contributions of research and innovation and stakeholder involvement in ethical governance of R&I can be of relevance.
- Awareness-raising on institutional documents through training and workshops. Here the guidelines on potential societal contributions of research and innovation, especially the part on the role of the Code of Ethics and Good Practice and how to communicate its contents in interactive workshops and other formats can help.

Conducting regular training programs and motivating stakeholders at all career levels to engage points to the issue of sustainability, which is of major relevance to projects such as the ETHNA System. The fact that implementing organizations report positively on the widening of their networks, initiated by the stakeholder engagement activities, especially the six steps of systematic stakeholder engagement, is deemed a success. As far as the sustainability and long-term implementation of the ETHNA System is concerned, the implementing organizations assume three main factors, all of which require the continuous involvement of relevant stakeholders. Firstly, a strong focus on communicating the content is needed, e.g., through training and other stakeholder activities, especially for early career researchers. In addition, in-depth knowledge of the organizational setting in which the ETHNA System is to be implemented is necessary, i.e., existing structures and conditions, values, needs and expectations of those directly or indirectly affected by the implementation, and the broader research environment. Finally, continuous updates and adjustments, particularly by incorporating feedback from implementers, are crucial to making stakeholder engagement a success [30].

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission.

References

1. Braun, R., Starkbaum, J.: Stakeholders in research and innovation: towards responsible governance. In: Blok, V. (ed.): Putting Responsible Research and Innovation into Practice. A Multi-Stakeholder Approach. Library of Ethics and Applied Philosophy **40**, pp. 229–248 (2023), https://doi.org/10.1007/978-3-031-14710-4
2. Owen, R., Macnaghten, P., Stilgoe, J.: Responsible research and innovation: from science in society to science for society, with society. Sci. Public Policy **39**(6), 751–760 (2012). https://doi.org/10.1093/scipol/scs093
3. Voegtlin, C., Scherer, A.G.: Responsible innovation and the innovation of responsibility: governing sustainable development in a globalized world. J. Bus. Ethics **143**(2), 227–243 (2015). https://doi.org/10.1007/s10551-015-2769-z

4. European Commission (EC): Rome Declaration on Responsible Research and Innovation in Europe (2014). https://www.sis-rri-conference.eu/wp-content/uploads/2014/12/RomeDeclaration_Final.pdf
5. Von Schomberg, R.: A vision of responsible research and innovation. In: Owen, R., Bessant, J., Heintz, M. (eds.) Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society, pp. 51–74. Wiley, Hoboken (2013)
6. González-Esteban, E.: The ETHNA System and Support Tools. In: Ethics and Responsible Research and Innovation in Practice, Springer, Cham (2023)
7. Andersson, E., et al.: What the Future Holds for Societal Engagement. Future Engagement Report. ENGAGE2020 D4.2 (2015). http://engage2020.eu/media/D4.2-Future-Engagement.pdf
8. Forsberg, E.-M., Gerber, A., Carson S.G.: Including RRI in the development and implementation of Horizon Europe (position paper) (2020). https://blog.rri-tools.eu/-/including-rri-in-the-development-and-implementation-of-horizon-europe
9. Kujala, J., Sachs, S., Leinonen, H., Heikkinen, A., Laude, D.: Stakeholder engagement: past, present, and future. Bus. Soc. **61**(5), 1136–1196 (2022). https://doi.org/10.1177/000765032 11066595
10. Blok, V.: From participation to interruption: towers and ethics of stakeholder engagement, participation and partnership in corporate social responsibility and responsible innovation. In: Von Schomberg, R., Hankins, J. (eds.) International Handbook on Responsible Innovation. A Global Resource, pp. 243–258. Edward Elgar, Cheltenham (2019), https://doi.org/10.4337/9781784718862.00024
11. Brand, T., Blok, V.: Responsible innovation in business: a critical reflection on deliberative engagement as a central governance mechanism. J. Res. Innov. **6**(1), 4–24 (2019). https://doi.org/10.1080/23299460.2019.1575681
12. Timmermans, J., Blok, V., Braun, R., Wesselink, R., Øjvind Nielsen, R.: Social labs as an inclusive methodology to implement and study social change: the case of responsible research and innovation. J. Res. Innov. **7**(3), 410–426 (2020). https://doi.org/10.1080/23299460.2020.1787751
13. Loureiro, P., Conceicao, C.: Emerging patterns in the academic literature on responsible research and innovation. Technol. Soc. **58**, 101-148 (2019). https://doi.org/10.1016/j.techsoc.2019.101148
14. Greenwood, M.: Stakeholder engagement: beyond the myth of corporate responsibility. J. Bus. Ethics **74**, 315–327 (2007). https://doi.org/10.1007/s10551-007-9509-y
15. Repo, P., Matschoss, K.: Considering expert takeovers in citizen involvement processes. J. Res. Innov. **6**(2), 119–142 (2019). https://doi.org/10.1080/23299460.2019.1568145
16. Häberlein, L., Mönig, J.M., Hövel, P.: Mapping stakeholders and scoping involvement. A guide for HEFRCs. ETHNA System Project – Deliverable 3.1 (2021a). https://doi.org/10.5281/zenodo.6088255
17. Häberlein, L., Mönig, J.M., Hövel, P.: Gauging the potential societal contributions of research and innovation – a guide for HEFRCs. ETHNA System Project – Deliverable 3.2 (2021b). https://doi.org/10.5281/zenodo.6088407
18. Häberlein, L., Mönig, J.M., Hövel, P.: Stakeholder involvement in ethical governance of R&I. A guide for HEFRCs. ETHNA System Project – Deliverable 3.3 (2021c). https://doi.org/10.5281/zenodo.6088546
19. Schütz, F., Heidingsfelder M.L, Schraudner, M.: Co-shaping the future in quadruple helix innovation systems: uncovering public preferences toward participatory research and innovation. J. Design, Economics, Innov. **5**(2), (2019). https://doi.org/10.1016/j.sheji.2019.04.002
20. Creighton, J.L.: The Public Participation Handbook: Making Better Decisions Through Citizen Involvement. Wiley, Hoboken (2005)

21. Moan, M.H., et al.: ETHNA system: literature review and state of the art description Mapping examples of good governance of research innovation (R&I) related to responsible research and innovation (RRI), in Higher Education, Funding and Research Organisations (HEFRs) in Europe (2021). https://doi.org/10.5281/zenodo.5814680
22. Lindemann, T.: Report on a Strategy to Engage Stakeholders, ROSiE D3.1 (2021)
23. Ravn, T., Mejlgaard, N.: Public Engagement Innovations – Catalogue of PE initiatives. PE2020 D1.2 (2015). https://tuhat.helsinki.fi/ws/portalfiles/portal/156529598/Public_En-gagement_Innovations_H2020.pdf
24. Rask, M., Matschoss, K., Kaarakainen, M.: Final Progress Report. PE2020 D6.1 (2017). https://cordis.europa.eu/docs/results/611/611826/final1-d6-1-final-progress-report-final.pdf
25. Durham, E., Baker, H., Smith, M., Moore, E., Morgan, V.: The BiodivERsA Stakeholder Engagement Handbook. BiodivERsA, Paris (2014). https://www.biodiversa.org/706/download
26. The Engage2020 Consortium: Engage2020. Science, Society and Engagement. An e-anthology. ENGAGE2020 D2.2 (2015). http://engage2020.eu/media/Engage2020_withVideo.pdf
27. D'Angelo, C., et al.: The use of public engagement for technological innovation. Literature review and case studies. BEIS Research Paper Number 2021/003 (2021)
28. Shelley-Egan, C., et al.: Report (handbook) of Participatory Processes, SATORI D 2.1 (2014). https://satoriproject.eu/media/D2.1_Report-handbook-of-participatory-processes_FINAL1.pdf
29. González-Esteban, E., Feenstra, R., Calvo, P., Sanahuja, R., Fernández-Beltrán, F., García-Marzá, D.: Final Concept of the ETHNA System. ETHNA System project – Deliverable 4.2 (2022). https://ethnasystem.eu/wp-content/uploads/2022/07/D4.2_ETHNA_2022_guide_220210_incl_toolbox_neu.pdf
30. Alves, E.: Report on the ETHNA System Implementation Analysis. ETHNA System Project – Deliverable 5.4 (2022a). https://doi.org/10.5281/zenodo.7501007
31. Alves, E: Report on the Difficulties Found in the Implementation Process. ETHNA System Project – Deliverable 5.5 (2022b). https://doi.org/10.5281/zenodo.7501022

Experiences and Lessons Learned

Evaluation of RRI Institutionalisation Endeavours: Specificities, Drivers, Barriers, and Good Practices Based on a Multi-stakeholder Consultation and Living Lab Experiences

Gábor Szüdi[1](✉), Dietmar Lampert[1], Marko Hajdinjak[2], Desislava Asenova[2], Elsa Alves[3], and Maya Vestergaard Bidstrup[3]

[1] Centre for Social Innovation, Linke Wienzeile 246, 1150 Vienna, Austria
SZUDI@ZSI.AT

[2] Applied Research and Communications Fund, Alexander Zhendov Street 5, 1113 Sofia, Bulgaria

[3] The Danish Board of Technology, Arnold Nielsen Boulevard 68E, 2650 Hvidovre, Denmark

Abstract. The aim of this study is to provide research performing organisations and research funding organisations (RPOs and RFOs) with practical advice on how to engage in an effective RRI institutionalisation. Therefore, we first looked at the most relevant drivers, challenges, and the most beneficial good practices potentially affecting RRI institutionalisation within RPOs and RFOs across Europe through a multi-step, multi-stakeholder consultation approach. The broad set of drivers, barriers and good practices identified at the consultation was methodologically divided into structural, cultural and interchange-related aspects. These aspects can theoretically exercise a positive or negative impact on the RRI institutionalisation, and their validity was tested in Living Labs by six organisations. By categorising these six implementers in terms of RRI readiness we were able to identify key factors and describe specific organisational circumstances conducive to a successful adoption and use of RRI principles and practices for three organisational types of RPOs/RFOs.

Keywords: RRI · Living Lab · RRI institutionalisation · drivers · barriers · good practices

1 Introduction

The main objective of the ETHNA project funded by the European Commission under the Horizon 2020 programme was to test the implementation potential of a novel governance, management and procedural system for Responsible Research and Innovation (RRI). This RRI governance system[1] aimed to design and promote sustainable research

[1] Titled ETHNA System, a concept that was finalised under the guidance of the project coordinator Jaume I University (UJI), more details on the specific content of the system can be found in [2].

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 57–72, 2023.
https://doi.org/10.1007/978-3-031-33177-0_4

and innovation practices respecting state-of-the-art ethical standards and other policies facilitating the adoption of RRI principles. The purpose was to develop a flexible system which can be adapted at as many research-performing organisations and research funding organisations (RPOs and RFOs) as possible.

Therefore, the envisaged concept defined only the broader competences, functions and structure of an 'ideal' RRI governance system with a proposed set of ethical management tools, and included good practices of various ethical management methods. The goal was to provide a truly practical guide of RRI institutionalisation, tailored to the vastly different circumstances, needs and requirements of RPOs and RFOs across Europe.

We define institutionalisation in practical terms, i.e., as the process of embedding RRI practices in a RPO or RFO in order to achieve some desirable organisational change where, at the very end of the process, RRI becomes an integral part of the organisation's identity, structure and culture which is not dependent on the effort of specific people [1].

In order to enhance the practical usability of the conceptual RRI governance system, the most relevant drivers behind RRI institutionalisation and the most serious barriers hindering the adoption or successful use of RRI within RPOs and RFOs across Europe had to be explored and better understood. This was achieved by a multi-stakeholder consultation process consisting of several steps, also striving to identify good practices of RRI institutionalisation.

The identification and analysis of these drivers, barriers and good practices took place before the pilot implementation (test phase) of the RRI governance system with the aim of providing future pilot implementers with practical advice on how to mitigate the risk of potential challenges that could threaten the success of the RRI institutionalisation process (*pre-pilot evaluation*).

The pilot implementation was carried out within six Living Labs. A Living Lab is a social, physical and/or mental space for social innovation and experimentation where different actors – both internal and external to the implementing organisations – come together with the objective to deal with a complex problem, and to come up with new and better solutions through dialogue, testing and reflecting. In our specific case internal and external stakeholders in six RPOs and RFOs pulled their resources to co-create, experiment with and evaluate the feasibility of the conceptual ideas and implementation steps embedded in the envisaged RRI governance system.

The Living Lab test phase ended with an in-depth evaluation of the process and outcomes, highlighting also the necessary conditions required to initiate, manage and implement RRI institutionalisation in different RPOs and RFOs. This evaluation also considered the specific drivers, barriers and good practices of different Living Lab contexts (*post-pilot evaluation*).

In practical terms the *pre-pilot evaluation* provided us with a broad set of potential drivers of and barriers to RRI institutionalisation that might be valid for different RPOs and RFOs. This broad set of factors were subsequently validated during the concrete implementation steps of the Living Labs, i.e. by using a small sample of diverse RPOs and RFOs a *post-pilot* evaluation was carried out where, on the basis of an original methodological framework, the *pre-pilot evaluation* results were placed into a broader context.

The contextualisation involved the understanding of the specific organisational circumstances and characteristics related to the most relevant identified drivers of and barriers to RRI institutionalisation.

The aim of this study is to use this contextualised data in order to provide RPOs and RFOs with practical lessons learnt on the incentive factors and challenges of RRI institutionalisation, coupled with specific measures on how to overcome such hindrances in the most effective way. We strive to give recommendations that can be useful for different types of RPOs and RFOs looking into the possibilities of adopting RRI principles and practices.

2 The Multi-stakeholder Consultation - Process and Outcomes

As detailed in the Introduction, a draft concept for a RRI governance system was elaborated in the course of 2021, providing a first interpretation of the proposed competences, functions, structure and ethical management tools of the system. However, a more nuanced, evidence-based and practical guidance on how to foster RRI institutionalisation for the foreseen target group of RPOs and RFOs was missing from this draft.

In order to make the concept adaptable to different organisational circumstances, key organisational factors advancing or hindering the institutionalisation of RRI within European RPOs and RFOs had to be identified and evaluated. For this purpose, a multi-stakeholder consultation was set up and implemented, gathering detailed information from a broad range of stakeholders across Europe on the status of RRI institutionalisation, the relevant drivers and barriers concerning the adoption of RRI principles and practices, as well as on the related good practices and implemented measures.

2.1 The Process of the Multi-stakeholder Consultation

The multi-stakeholder consultation was conducted between January and October 2021 before the Living Lab pilot implementation, consisting of a preliminary phase and three subsequent phases, i.e., *interviews, workshops* and a *global survey.* Each stage focused on better understanding the relevant drivers, barriers and good practices concerning the adoption of RRI measures in RPOs and RFOs (Fig. 1).

Fig. 1. The Phases of the Multi-stakeholder Consultation and Living Lab Experience. Source: ZSI's own rendition

The *preliminary phase* used a questionnaire to be answered until the end of January 2021 only by the six Living Lab implementers to assess their currently available and missing RRI tools, initiatives and aspects, and their future plans to deal with missing RRI dimensions. Importantly, the questionnaire helped evaluate the interdependencies between organisation types, research area(s), organisational structure and RRI

uptake, which contributed to developing a future self-assessment method for 'RRI readiness' (RRI institutionalisation quadrants) and outlining the focus of the interviews and workshops.

In the *first consultation phase* 25 semi-structured interviews were conducted between April and June 2021 with mid- or high-level representatives of European RPOs and RFOs actively engaged in RRI governance or specific RRI keys. People familiar with key concepts of RRI and organisational or national adoption of RRI principles and keys were interviewed about their expertise and understanding of RRI, the organisational drivers and barriers of RRI institutionalisation, as well as good-practice examples or ideas.

We aimed to involve stakeholders from across as many countries, RRI keys and organisation types as possible, also striving to conduct interviews with persons working in research (or innovation) performing businesses and civil society organisations. Therefore interviews were made with representatives of the following ten countries: Austria, Bulgaria, Denmark, Estonia, Hungary, Norway, Portugal, Serbia, Slovakia, Spain. Altogether seven RFOs and 18 RPOs (including five businesses and two civil society organisations) were interviewed.

Their opinion was considered important as an input for the Living Labs that by definition aim to foster interactions of four stakeholder groups within the knowledge economy, namely academia, business, policy-makers and civil society [3–5].

Based on the responses of the interviews, five online workshops were organised per RRI keys with Living Lab implementers and external expert stakeholders (participant number ranging from 11 to 17) as a *second consultation phase* between July and September 2021. The workshop participants discussed in more detail the factors driving or hindering RRI institutionalisation in RPOs and RFOs.

The workshops aimed to obtain relevant observations and develop recommendations for the Living Labs with respect to the relevant drivers, incentives, barriers and organisational strategies and practices for RRI governance. Their main outcome was the ranking of drivers and barriers of RRI institutionalisation per RRI keys and organisational types, which was verified by an online survey in the *third consultation phase.*

The online survey was sent in October 2021 to a broad group (10.000 +) of potentially relevant expert stakeholders identified through a Web of Science database search with the aim of assessing the relevance of the identified drivers, barriers and good practices of RRI institutionalisation. Altogether 888 responses were received from 69 countries with a balanced gender representation, involving the opinion of mostly senior experts (55% having more than 15 years of experience)[2].

After filling out general demographic and organisational information, the respondents rated the perceived relevance of the RRI incentives, barriers and good practices that were the highest ranked at the end of the second consultation phase (using a Likert scale of 1–10). Thus, the survey confirmed the results of a smaller, unavoidably biased sample on a global scale with the support of RRI experts. These results are summarised in the next Section.

[2] See results at: https://zenodo.org/record/6616553#.ZBmaKXbMI2x.

2.2 Lessons Learnt from the Multi-stakeholder Consultation

To present our findings, we chose to use as a basis the theoretical framework developed by [6] in the Horizon 2020 project RRI-Practice that was further refined in [7]. This framework was already utilised in the workshop phase when relevant RRI drivers and barriers, as well as potential organisational actions were divided into structural, cultural, and interchange perspectives. All three perspectives focus on different aspects within organisations (RPOs and RFOs) that might contribute to or hinder the institutionalisation of RRI or its keys.

The first perspective focuses on regulative and normative aspects that structure and standardise organisational behaviour. Structural aspects include formalised roles and positions, mandates, responsibilities, decision-making structures in the organisation, namely persons or units specifically tasked with RRI-related duties in RPOs and RFOs such as ethics boards or research integrity offers, and the related formal and informal documents, namely concepts, norms, standards, procedures and strategies, such as code of ethics, gender equality plans open science policy guidelines.

The second perspective focuses on cultural aspects and, consequently, deals with the informal and tacit organisational structures that influence the adoption of RRI standards and practices. These structures might explain the difference and the interconnected relation between policy goals and practical behaviour within organisations. Cultural aspects include organisational cultures, values, and identities, i.e. for the purposes of this article, the perceptions about RRI.

Third, the interchange aspects are based on the observation that organisations are not only influenced by their structure and culture but also by their interactions with other organisations in their broader environment. Thus, these aspects focus on drivers or barriers stemming out of but connected to the organisation, such as impacts of the broader policy landscape or research culture [7].

RRI governance was regarded throughout the consultation as a horizontal, cross-cutting dimension, which means it had gained an elevated role which was then reflected by Living Labs through their focus on the whole RRI governance structure. Table 1 summarises the aspects deemed most relevant by stakeholders in terms of adoption and successful use of RRI in their respective organisation. A more detailed analysis at the level of specific RRI keys may be of interest but is outside the scope of this article.

As regards the potential drivers for RRI, the stakeholders identified two key structural factors for a successful RRI institutionalisation process, namely a supportive (higher and mid-level) management, and support structures and practices that can take many forms within an organisation, including pilot programmes dedicated to certain RRI keys, organisational units or infrastructure dealing with specific RRI keys. A management keen on furthering RRI also should strive to adopt organisational mandates, regulations or strategies that prescribe concrete policy goals aimed at fostering RRI.

Such formalised drivers are more effective in an organisational environment aligned through values or identity with the overall concept and specific aspects of RRI. RRI should become part of the culture within an organisation when research activities at all stages take into consideration RRI keys, such as ethics and research integrity. Organisational ‘facilitators’, i.e. persons or units already engaged in RRI and willing to share

Table 1. Structural, cultural and interchange aspects of RRI governance.

	Structural aspects	Cultural aspects	Interchange aspects
Potential drivers for RRI	- Organisational mandates, regulations, strategies - Management keen on furthering RRI - Support structures and practices, incl. Dedicated pilot programmes	- Organisational values and identity - Organisational 'facilitators'	- Requirements or expectations from funding bodies - Adherence to national or EU standards or normative laws
Potential barriers to RRI	- Lack of resources (human, financial, time, etc.) - Lack of institutional support structures and practices - Lack of support from management	- Lack of awareness - Lack of understanding - Lack of motivation - Lack of institutional values, standards and visions	- Lack of or confusing policies, strategies or mandates
Potential organisational measures (good practices)	- Knowledge pooling within the organisation - Performance metrics - Trainings and education - Awards	Reflection spaces Engagement with external stakeholders Practical guides	- Alignment with external standards and funding requirements

Source: authors' categorisation based on the methodological framework by [7], taken from [8].

their knowledge and skills with others, are perceived crucial by stakeholders to embed RRI both in the culture and mandate of a given organisation.

If internal incentives are not strong enough to facilitate the uptake of RRI, external drivers can gain in importance. More and more often research funders by mandate aim to 'nudge' RPOs to adopt RRI aspects through assessment criteria (compulsory gender equality plans), monitoring requirements or other policies. In addition, a proactive management might seek to voluntarily align organisational practices with national standards or international benchmarks in various RRI keys.

Nevertheless, potential barriers to RRI might counteract with the impact of these drivers. The majority of the stakeholders consider the lack of resources as the most influential barrier. This simply means that there are not enough employees to deal with RRI or researchers under constant pressure have no time to engage in RRI. The adequate funding is also often missing to compensate for such extra work.

This manifests in a lack of or underdeveloped support structures and practices when RPOs are not able to focus on RRI aspects in addition to their 'core' research activities. The root problem might be the lack of engagement from a managerial level but such a

state might merely develop due to financial and time constraints in spite of the better efforts of the leadership.

Such management efforts might be also hampered by a lack of awareness or understanding of RRI and its keys. RRI as a concept is still not completely clear for many members of the research community, especially if they do not deal with EC research and innovation policies or projects. In many cases some RRI keys are present in the organisation but not perceived as part of the RRI concept. Knowledge is fragmented between units, teams and researchers resulting in a lack of explicit support for RRI in organisational values, standards and long-term vision.

The fragmentation might be exacerbated by confusing RRI-related policies, strategies or mandates of policy-makers or research funders in different research fields or countries that sometimes contradict each other or give no clear guidance. Such a state of affairs might demotivate even the most dedicated and diligent researchers.

The stakeholders described several organisational measures to combat such demotivation trends and foster RRI institutionalisation. Even though the aim of aligning the organisation – or its dedicated units – with the most important external standards and funding requirements might seem inconceivable in the short term, more practical steps at lower levels can be started to foster RRI uptake.

In order the decrease fragmentation of knowledge on RRI, the organisation should carry out knowledge pooling measures, e.g., by using internal information repositories to check what RRI-related documents are already available within the institution, or by implementing a work environment survey to gather data on the readily available institutional knowledge on RRI-related principles and measures.

Reflection spaces organised in formal or informal settings, such as dialogue sessions or workshops, with the aim of knowledge pooling might contribute to understanding how to build RRI concepts and practices into research activities. In addition to internal knowledge sharing, networking with external stakeholders within and outside of other RPOs might also facilitate knowledge transfer with regard to the identification of common problems and solutions in RRI issues.

The increased and less fragmented knowledge base should be compiled and disseminated in practical guides that also aim to reduce the lack of conceptual and terminology clarity of RRI. In order to raise awareness of RRI issues, motivate people to get engaged and gain top-down support, the benefits of RRI should be explained in a clear and understandable manner.

The sustainability of such measures can be ensured by a couple of good practices frequently mentioned by the stakeholders, such as practical training or more formal education opportunities offered to researchers and other employees, institutional or departmental awards or other recognitions given to individuals or teams raising awareness of or tackling RRI-related issues. Finally, revised performance metrics evaluating the progress of RRI uptake at organisational or departmental level would be crucial for ensuring sustainable positive changes.

Evaluating organisations through the lens of the aforementioned structural, cultural and interchange aspects provide a broad overview on the possible uses of RRI in RPOs and RFOs but their practical relevance could only be validated through the actual Living

Lab implementation also taking into consideration the more theoretical findings of the multi-stakeholder consultation with the involvement of external stakeholders.

The above summarised results of the stakeholder consultation supported the Living Lab implementation in two significant respects. First, they contributed to the development of a practical toolbox to support implementation approaches and to provide Living Lab implementers with options to spark ideas on the potential challenges and ways to overcome those (while helping understand what drives the institutionalisation of RRI in the first place). Second, the results laid the groundwork for categorising the potential usefulness and adaptability of the concept in different organisational settings, which will be described in-depth in the next subsection.

2.3 The RRI Institutionalisation Quadrants

We found two critical factors to assess the given organisation's readiness for RRI institutionalisation through the analysis of the initial results gained from the different phases of the multi-stakeholder consultation. Notwithstanding the fact that there are other relevant institutional factors influencing the chances of a successful adoption and use of RRI by RPOs and RFOs – such as among others size, country, research area, funding sources – two dimensions can be considered as essential for the institutionalisation of RRI: the leadership and the base (Fig. 2).

First, the leadership represents the commitment of the higher or mid-level management in terms of top-down support received by organisational stakeholders for RRI institutionalisation. The leadership might provide such top-down support in various ways ranging from a passive commitment to a more active involvement, such as the heightened awareness level of certain RRI-related issues, the prevalence of a clear vision on RRI, the willingness to adapt conditions or allocate the required – human or financial – resources for a better RRI institutionalisation.

Second, the base means the strength of the organisational structures in the support of RRI institutionalisation that are launched and managed at various (lower) levels of the organisation. As opposed to the leadership dimension, these support measures are mostly started from bottom-up, enabled by the values, awareness, skills and knowledge of the research staff and other organisational stakeholders.

Applying leadership to the y-axis and the base to the x-axis results in a two-dimensional system with four quadrants that can be characterised as follows:

A *strong leadership but weak base* (top left quadrant) means that initiatives to drive the institutionalisation of RRI may already be under way but, generally, have not borne fruit yet, i. e. RRI norms and practices have not been broadly adopted by the base yet. However, the leadership is strong in terms of providing guidance on designing and implementing relevant activities. This guidance could be reflected in an increased awareness, sense of urgency, willingness, clarity of vision, leadership skills, resources, etc.

A *strong base but weak leadership* (bottom right quadrant) means that RRI initiatives can already be found in the organisation however the leadership is weak in terms of RRI institutionalisation. This weakness might entail that the management might not have heard about such initiatives, might not care, or might think that these initiatives are too specific or small to be transferred, scaled up, or adopted by all organisational units. In

Fig. 2. The RRI Institutionalisation Quadrants – Leadership and Base. Source: ZSI's own rendition, taken from [9]

theory, the organisational focus should first lie on spreading RRI norms and practices locally, on building showcases, and on connecting to similar efforts – both internal and external ones – to build a critical mass and reach and engage the leadership.

A *strong leadership and strong base* (top right quadrant) means that both leadership and base are aligned in terms of RRI institutionalisation needs and efforts. Such organisations show advanced levels of RRI institutionalisation and might have a long tradition of reflecting and adjusting their research practices, of reacting to external normative efforts (e.g., the adoption of standards), and of building institutional support structures and mechanisms. The guidance in this case will focus on giving impulses with the aim of further refining the institutionalisation efforts and adopt an anticipatory perspective in terms of future developments [9].

The devised RRI governance system was designed to work for all quadrants, with the exception of the lower left quadrant, i.e., weak leadership in combination with a weak base. A major prerequisite for the RRI institutionalisation is that at least one dimension needs to be at least somewhat strong, otherwise there is nothing to build on. For a successful institutionalisation process, both dimensions need to become strong in the long run.

This RRI Institutionalisation Quadrants helped the Living Labs to gauge their strength and determine their level of commitment to the options provided by the RRI governance concept.

The next Section will give an overview on the Living Lab implementation, its evaluation process and the most relevant evaluation results that are summarised with the help of the above-described conceptual framework.

3 The Living Lab Implementation – Process and Evaluation

In order to test the practical applicability of the finalised RRI governance concept in different RPOs and RFOs, it was planned and implemented at six organisations based in five different European countries.

The following organisations participated in the Living Lab experiment: University Jaume I (UJI) [10], a large public university from Spain; the Norwegian University of Science and Technology (NTNU), the largest public university in the country; the Education and Youth Board (Harno), a funding agency from Estonia; the Science, Technology and Business Park (ESPAITEC), a Spanish technology park; the Institute for the Development of New Technologies (UNINOVA) [11], a multidisciplinary, independent, and non-profit research institute from Portugal, and Applied Research and Communications Fund (ARC Fund) [12], a non-profit research and innovation policy institute from Bulgaria..

The implementation process followed the Living Lab methodology, and was divided into six stages (planning; construction; consultation; refinement; testing; review) lasting approximately one year (November 2021 – October 2022), with some institutions ultimately experiencing delays in certain process stages.

Since not much time has passed since the official end of the Living Lab experiment at the time of writing this study[3] it is not yet feasible to assess the scope of the institutional changes induced by the implementation process as the actual impact will only become tangible in a longer timeframe. However, it is already possible to draw conclusions at a more practical level. Therefore, the evaluation process focused on the most common drivers and barriers of implementation, highlighting concrete actions and good practices emerging from the process, as well as outlining necessary conditions for supporting the organisational uptake of RRI in the six implementing organisations.

The evaluation took place between September and November 2022 and contained the following steps:

1. DBT organised two rounds of two online, 3-h participatory evaluation workshops in September and October 2022. The first round of workshops was held with representatives of the project partners participating in Living Labs, while the second round of workshops involved internal and external stakeholders supporting the Living Lab implementation.[4].

   The workshops were organised as semi-structured events focused on the added value of the RRI institutionalisation process in general, as well as the particular experiences encountered in the actual implementation process. The workshops employed a diversity of exchange formats to support mutual learning and feedback gathering as needed. This structure provided the participants with great flexibility to engage in

[3] December 2022 - March 2023.

[4] More information on the format and results of the participatory evaluative workshops is available here: Alves, Elsa (2022). D5.4 Report on the ETHNA System Implementation Analysis & Alves, Elsa (2022). D5.5 Report on the difficulties found in the implementation processes. https://ethnasystem.eu/wp-content/uploads/2023/01/5.5-Report-collecting-the-difficulties-found-in-the-implementation-processes-final_181222.pdf

dialogue and generate collective reflections on the insights and lessons learned from the implementation process.

The purpose of the workshops was to create a common space for the Living Lab stakeholders to critically reflect on their experiences with the implementation and directly share the matter-of-fact assessment of their hands-on experiences. As an end result, the workshops contributed to the elaboration of more specific evaluation questions about the methodology and process of the implementation.

2. The in-depth evaluation questions were asked from the so-called Lab Managers, the key people responsible for the planning, coordination and facilitation of the Living Lab implementation process. Their responsibility ranged from implementing and monitoring all stages of the process through recruiting, engaging and supporting all relevant internal and external stakeholders to communicating and reporting to the organisational and project management.

   Late October and early November 2022 the Lab Managers answered the questions in the form of an online self-evaluation questionnaire developed by ARC Fund. They had to give short but concise answer to a variety of questions introducing their organisation, explaining the reasons for their commitment to adopt institutional changes, and going into detail about the actual measures undertaken, also highlighting the participating internal and external stakeholders, as well as the barriers, drivers, good practices and potential sustainability of the induced changes.
3. Building on the responses of the Lab Managers, a 2-day workshop dedicated for knowledge and experience transfer was organised by ARC Fund at the end of November 2022. On the first day the lessons and experiences of the six implementation cases were discussed in detail with the involvement of external experts, and on the second day a final evaluative workshop was organised under the guidance of ZSI on the emerging challenges and potential sustainable outcomes of the Living Labs.

   This workshop was the first opportunity to bring together, in one physical location, all Living Lab implementers to discuss and rank the barriers to the implementation of the elaborated RRI governance system within organisations and the measures to react to them.

The work was done in break-out sessions with facilitators where one group included the Lab Managers, while two other break-out sessions involved the other internal and external Living Lab stakeholders to discuss the key enabling external factors (drivers) and the actions and strategies that could exercise an either positive or negative impact for RRI institutionalisation. After the end of the parallel break-out sessions the group work was discussed in a plenary setting involving all participants to draw conclusions and make recommendations.

3.1 Lessons Learnt from the Living Labs

By resorting to the drawn-up methodological framework dividing the six implementer organisations to three categories (quadrants) in respect of readiness for RRI institutionalisation we are able to put the evaluation results to a broader context and describe the specific organisational circumstances and characteristics relating to the identified drivers and barriers.

Each organisation has made a self-assessment in terms of the RRI institutionalisation quadrants at the second phase of the Living Lab evaluation (online questionnaire). UJI and Harno have indicated to belong to the 'strong leadership, strong base' category. UNINOVA and ESPAITEC were classified to the 'strong leadership, weak base' category. ARC Fund confirmed to have a strong base but a weak leadership. NTNU could not decide on a specific category but upon further investigation we have added the institution to the same category (For a longer description of Harno, UNINOVA, and UJI cases see, respectively, chapters 5, 6 and 7 in this same volume).

We add this additional layer of organisational diversity to the already established three types of RRI governance perspectives, i.e., the structural, cultural and interchange aspects to summarise our findings as follows:

In the 'ideal' scenario when there is both a supportive management and already established support structures and practices for RRI institutionalisation, many barriers have already been removed. Most importantly, the dedication of the management means a favourable position in getting the necessary extra funding and experts needed for RRI institutionalisation within the organisation. Managerial support and the previous good experience gained with RRI practices also increase the chances of a successful cooperation of internal and external stakeholders across several disciplines.

While structural barriers have lost their relevance in this case, cultural barriers might still need to be overcome: the theoretical support expressed by stakeholders should be turned into an active commitment with concrete contributions. In addition to a heavy workload, this reluctance might stem from a lack of knowledge and understanding of the complex RRI concept sometimes perceived as mandated by external parties.

Cultural drivers might help counteract such attitudes, aiming to embed RRI into organisational (soft) values and identity. A participatory and collaborative process was used to set up a truly open debate space to discuss how to achieve this goal. Such a process could benefit from a neutral organisational facilitator and definitely should involve external stakeholders.

In the concrete Living Lab implementation cases such a process consisted of an initial consultation and interviews about the knowledge of various RRI keys, internal working groups meetings, bilateral meetings and workshops with external stakeholders, also seeking synergies with broader initiatives of interests within the organisation.

The successful implementation of such open participatory processes is in itself a success but the Living Labs managed to adopt new codes of ethics and good practices in this short timeframe. Particularly important was the addition of a glossary of complex RRI concepts into the code drafted by UJI, which proved to be very useful for the interested research community. Such novel outcomes contribute to the dedicated objective of these Living Lab implementers to strengthen their frontrunner position in RRI and be recognised for this level of excellence (a cultural driver).

Living Labs with a committed (higher-level) management but a weak base were in a special situation because they are not 'traditional' RPOs: UNINOVA's researchers are employees of other institutions and therefore subject to the host organisations' RRI governance systems (in this sense UNINOVA can be seen as a federated ecosystem of

researchers), while ESPAITEC is a science and technology park acting as a facilitator among all the innovation ecosystem agents (start-ups, spin-offs, entrepreneurs and university researchers).

Hence managers aimed to focus on complementary RRI aspects perceived as the most important or the most feasible to implement (e.g., gender in case of ESPAITEC). The managers wanted to promote excellent research and innovation practices to 'change the organisational culture' but were also motivated by interchange-related drivers, e.g., to comply with the contractual obligations towards the national research funding agency in terms of RRI.

As a first step, a knowledge pooling exercise was conducted to identify the achievable goals and priorities by recognising the weaknesses to overcome, the already available RRI knowledge and skills, and the best methods to adopt elements of RRI governance in a sustainable way.

Similar to Living Labs with a strong base and leadership, the scarce time available for researchers to spend on RRI issues proved a major barrier. To increase the awareness, understanding and motivation of researchers towards RRI institutionalisation, organisational 'facilitators' (a small but dedicated RRI team or external experts) planned a participatory process with as few formalities as possible to consult, refine and adopt practical RRI documents that also adapted RRI jargon to institutional reality. An important aspect is that external stakeholders were involved in this process through working sessions to offer valuable ideas, feedback and networking opportunities for an effective RRI institutionalisation.

The participatory process resulted in drafting key documents on various RRI aspects and in case of UNINOVA was complemented by actions with the aim to change organisational culture, such as a specific website section to raise awareness on RRI, or training sessions on RRI organised for young researchers and PhD students of the institute.

The implementation process progressed with the most difficulties in Living Labs where – opposite to the above-described scenario – the existing base was strong but leadership support was ambivalent or remained declarative in nature.

In both cases the base was strong: there were clear organisational mandates to conduct research in an ethical way, for the public benefit and support structures were already in place in the form of various documents and (advisory) bodies (even though scattered around in different departments or not explicitly referring to RRI). The researchers in both organisations were also quite well-versed in and motivated to deal with RRI or ethical issues, based on their disciplinary or project-related experience and professional interests. External drivers such as the requirement of funding programmes might have also played a facilitating role for RRI uptake.

Nevertheless, key structural barriers prevented these Living Labs from achieving tangible results in the relatively short implementation timeframe defined by the project. The usual culprit of lack of time and personnel available for RRI institutionalisation was worsened by the indecisive support and engagement rendered by (higher-level) senior managers. Thus, even the initial decision on the proper implementation level and planning caused unwanted delays. This was connected with the issue of size: NTNU is too big for a Living Lab and thus looked for a suitable department for implementation, while ARC

Fund is too small and has very different foci around three thematic programmes (but in the end used RRI as a common frame).

The ambivalent support did not help convince the relevant stakeholders of the benefits of RRI institutionalisation: there was a general feeling among some researchers that this is an externally mandated process which aims to discuss again topics that have already been discussed and/or do not need solutions. The size of implementers played again a role here in different ways: the small team of ARC Fund researchers felt a sense of 'fatigue' towards the topics already encountered several times in many RRI-projects and NTNU has already possessed of similar RRI governance structures but at an organisational (not departmental) level.

In short, initial structural barriers exacerbated cultural barriers in turn. To remedy the situation a participatory process for RRI institutionalisation started but progressed slowly or were stuck in key moments. This process in both organisations successfully managed to assess the RRI-related situation, and identify important RRI aspects worthy of further discussion or endorsement but concrete supporting documents or bodies have not been adopted yet. The main tool used was different types of internal reflection spaces, such as semi-structured interviews, workshops and focus groups, however the engagement of external stakeholders was deemed problematic.

4 Conclusions

The 'post-pilot' evaluation confirmed the validity of some of the key findings of the 'pre-pilot evaluation', meaning that certain incentive factors and barriers need to be considered in specific organisational circumstances to ensure a successful organisational adoption and use of RRI.

Strong leadership, i.e., the active engagement and support of the higher-level management seems to be the most significant driver without which a genuine RRI institutionalisation is doomed to fail or at least progress slowly. If other key structural drivers are available, e.g., support structures ('base') and adherent organisational mandates, the RRI institutionalisation can progress steadily with the aim of introducing substantial changes in all RRI keys.

The existence of such supporting structure also stemming from bottom-up initiatives are particularly important for sustainability because it is true that the implementation may start as a top-down approach (even forced by external requirements e.g. from funding bodies), but its long-term impact ultimately relies on the bottom-up approach guaranteeing appreciation and motivation of relevant stakeholders.

This bottom-up approach is manifested in the co-creation process which lies at the heart of Living Labs aiming to reform the 'business-as-usual' approach to research in organisations. Co-creation can be in itself a challenge; e.g. regarding the involvement of external stakeholders, with which more implementers struggled – but it is also a rewarding endeavour where internal and external stakeholders from many disciplines participate in enriching discussions in various reflection spaces to improve the quality and relevance of the achieved results.

Co-creation is time and resource-intensive which is another barrier to take care of early on. In order to plan for feasible results with the available resources, Living Labs

should start with the understanding of the broader (country) and local (organisational) context, i.e. available funding, personnel, time, prevailing and missing RRI aspects, the perspective and needs of stakeholders, preferably done by organisational 'facilitators' (proactive and committed employees with experience in RRI).

Living labs are context-bound, there are no 'one-size-fits-all' solutions but goals and actions should be aligned with national and institutional reality. The methodology developed in the project is considered to provide good ideas and inspiration but each organisation should develop its own path towards RRI-paved institutional change.

The results show that substantial RRI-related changes in such a short time frame could only be achieved by larger RPOs and RFOs that have already had a strong leadership and base (e.g., UJI, Harno). Implementers with no strong leadership but with a formidable base could only use this time for self-reflection and a better understanding of their situation. Implementers with no base but an engaged management could go one step further and look for complementarities, i.e., adopting RRI aspects perceived as the most important to the organisation, subsequently benefitting the use of other RRI keys.

While the long-term goal of all such Living Labs should be the change of culture, the impact of seemingly small-scale changes should not be underestimated. A shift in organisational culture might be achieved exactly by such actions, e.g., hands-on guides with an understandable terminology, awards given for considerable RRI-related achievements or practical training on various RRI aspects, changing the RRI-related attitudes and mindset of the next generations.

Such an incremental approach is not only beneficial due to the high variety in institutional settings but also because many barriers are interconnected and reinforcing. Living Labs should use a flexible and adaptable approach to find the right intervention points to tackle the RRI-related issues deemed most relevant by stakeholders with the appropriate measures available within the context-dependent conditions. Thus, the RRI institutionalisation process will not be seen as a 'straitjacket' but rather as an opportunity.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission. The authors would like to thank the reviewers for their comments.

References

1. Steen, M., Nauta, J., Gaasterland, R., Ogier, S.: Institutionalizing responsible research and innovation: case studies. In: Bitran, I., Conn, S., Huizingh, E., Kokshagina, O., Torkkeli, M., Tynnhammar, M. (eds.) Proceedings of the XXIX ISPIM Innovation Conference: Innovation, The Name of the Game (2018) ISPIM
2. González-Esteban, E.: The ETHNA system and support tools. In: Ethics and Responsible Research and Innovation in Practice, Springer, Cham (2023)
3. Evans, P., Schuurman, D., Ståhlbröst, A., Vervoort, K.: Living Lab Methodology Handbook. European Network of Living Labs (2017). https://issuu.com/enoll/docs/366265932-u4iot-LivingLabmethodology-handbook. Accessed 15 Mar 2023

4. Häberlein, L., Mönig, J.M., Hövel, P.: Mapping stakeholders and scoping involvement – a guide for HEFRCs. Project Deliverable, Ethics Governance System for RRI in Higher Education, Funding and Research Centres (2021). https://ethnasystem.eu/wp-content/uploads/2021/10/ETHNA_2021_d3.1-stakeholdermapping_2110011.pdf. Accessed 15 Mar 2023
5. Popa, E., Blok, V., Wesselink, R.: Social Labs for Quadruple Helix Collaborations – A manual for designing and implementing social Labs in collaboration between industry, academia, policy and citizen. Project Deliverable, RiConfigure (2018). http://riconfigure.eu/wp-content/uploads/2019/10/D01.2_Social-Lab-Methodology-Manual_v2.0.pdf. Accessed 15 Mar 2023
6. Forsberg, E.-M., Ladikas, M.: Analytic framework and case study protocol. Project Deliverable, RRI-Practice (2017). https://www.rri-practice.eu/knowledge-repository/publications-and-deliverables. Accessed 15 Mar 2023
7. Wittrock, C., Forsberg, E.-M., Pols, A., Macnaghten, P., Ludwig, D.: Introduction to RRI and the organisational study. In: Implementing Responsible Research and Innovation. SE, pp. 7–22. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-54286-3_2
8. Szüdi, G., Lampert, D.: Blueprint for institutional change to implement an effective RRI governance. Project Deliverable, Ethics Governance System for RRI in Higher Education, Funding and Research Centres (forthcoming). https://doi.org/10.5281/zenodo.7730729
9. González-Esteban, E., Feenstra, R., Calvo, P., Sanahuja, R., Fernández-Beltrán, F., García-Marzá, D.: Draft concept of the ETHNA System, Project Deliverable, Ethics Governance System for RRI in Higher Education, Funding and Research Centres (2021). https://ethnasystem.eu/wp-content/uploads/2022/02/D4.2_ETHNA_final-concept.pdf. Accessed 15 Mar 2023
10. Bernal-Sánchez, L., Feenstra, R.A.: Developing RRI and Research Ethics at University. In: Ethics and Responsible Research and Innovation in Practice, LNCS, volume 13875, Springer, Cham (2023)
11. Camarinha-Matos, L.M., Ferrada, F., Oliveira, A.I.: Implementing RRI in a research and innovation ecosystem. In: Ethics and Responsible Research and Innovation in Practice, LNCS, volume 13875, Springer, Cham (2023)
12. Hajdinjak, M.: Implementing RRI in a Non-Governmental Research Institute. In: Ethics and Responsible Research and Innovation in Practice, LNCS, volume 13875, Springer, Cham (2023)

Implementing RRI in a Non-governmental Research Institute

Marko Hajdinjak(✉)

Applied Research and Communications Fund (ARC Fund), Sofia, Bulgaria
marko.hajdinjak@online.bg

Abstract. This chapter critically examines the outcomes from the ETHNA System implementation process in a small non-governmental research organization. The chapter briefly presents the process, identifies the main barriers and drivers pertaining to the implementation, and highlights the good practices that could serve as inspiration or model for other similar organizations. The chapter concludes with recommendations for organizations similar to ARC Fund and interested in implementing the RRI-based governance structure for management and conduct of research processes.

Keywords: Ethics governance structure · socially relevant research · responsible research and innovation · small non-governmental research organization

1 Introduction

Established in 1991, the Applied Research and Communications Fund (ARC Fund) is a Bulgarian innovation policy and applied research institute, whose mission is to contribute to the development of a modern, knowledge-based society, exploiting the power of information technologies and innovation. As one of the leading not-for-profit research-performing organizations in Bulgaria, ARC Fund has been for years an active promoter of the Responsible Research and Innovation (RRI) framework in the country. Since 2015, the organization has participated in several H2020-funded projects focused on RRI, namely RRI-Practice,[1] TeRRItoria,[2] SUPER MoRRI,[3] and ETHNA System,[4] and is a coordinator of another such project (RRI-LEADERS).[5]

Not surprisingly, RRI has become an increasingly important topic in the organization, leading to the realization that a more formal uptake of RRI concept on all levels of management (senior management, executive management and operational management) and among the research staff would be needed. Such a development would also significantly enhance the legitimacy of the organization as the RRI champion in Bulgaria

[1] https://cordis.europa.eu/project/id/709637.
[2] https://cordis.europa.eu/project/id/824565.
[3] https://super-morri.eu/.
[4] https://ethnasystem.eu/.
[5] https://www.rri-leaders.eu/.

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 73–86, 2023.
https://doi.org/10.1007/978-3-031-33177-0_5

vis-à-vis the key stakeholders in the society. Participation in the ETHNA System project offered such an opportunity in the form of development and implementation of an ethics governance system.

Although the entire RRI framework is highly relevant for ARC Fund, certain RRI keys are especially important. Open access, ethics and gender equality (included in the larger frame of non-discrimination) have been ingrained into ARC Fund's practices and objectives practically since its establishment. In the early years of this century, public engagement has become a crucial aspect employed both for defining research goals and priorities, and for verifying the research outcomes.

ARC Fund's overall approach to research process and knowledge governance is based on several leading principles: being proactive in terms of considering likely impacts of project-related work; articulating possible risks and foreseeing appropriate mitigation actions; involving various stakeholders and general public in different stages of the research process; and continuously reflecting on the original purposes and motivations for performing socially relevant and useful research. These circumstances were assessed to represent a solid base for implementation of the ETHNA System at ARC Fund.

This chapter critically examines the outcomes from the ETHNA System implementation process in a small non-governmental research organization such as ARC Fund. The chapter identifies the main barriers and drivers that emerged from the process, highlights the good practices that could serve as an inspiration for other similar organizations, and draws some conclusions about the potential scope and form of institutional changes induced by the application of the ETHNA System. The chapter concludes with recommendations for organizations similar to ARC Fund and interested in implementing the RRI-based governance structure for management and conduct of research processes.

2 Meaning and Relevance of Responsible Research at ARC Fund

In order to enable a smooth and effective process of ETHNA System implementation, a thorough mapping of existing RRI-related resources was conducted. This included a review of existing internal documents such as the Statute of ARC Fund (2007), the Code of Conduct (2008), the Code of Ethics (2008), the Rules and Order for Performing Publicly Beneficial Activities (2008), the Child Protection Policy (2016) and the Rules for the processing and protection of personal data (2018). After the document review, eight interviews with representatives of the senior and middle-level management were conducted in the period March-April 2022. Nine staff members from various positions participated in a focus group, organized in June 2022.

Based on this internal research, the meaning and relevance of responsible research at ARC Fund has been assessed. The responsibility in organization's research activities can be framed by the following principles:

- Adherence to rigorous ethical principles in everyday research activities: ARC Fund's researchers are obliged to strictly respect intellectual property rights and copyright regulations; to be impartial and objective in research; to thoroughly check the information and data sources; to interpret data in an objective and unbiased way and consider different views; and to avoid misuse of research results (research ethics and integrity).

- Use of only those research methods, which comply with rigorous ethical norms and moral imperatives: Do-no-harm principle must be always respected – this includes but is not limited to the impacts on the health of people and on the environment (research ethics and integrity).
- Respect for other participants involved in research: Fellow researchers, stakeholders, funders and representatives of target societal groups involved in or affected by the research must always be treated with due respect and consideration (research ethics).
- Understanding the needs of society: In order to address societal challenges, researchers at ARC Fund need to align the goals and directions of their research endeavors with the broader societal needs (public engagement).
- Consulting the process and results of the research with representatives of broader public groups and stakeholders: Whenever possible or/and relevant, different stakeholders should be involved in various stages of the research – from defining research goals and priorities to presentation of the research results in a way that is understandable by the society at large, i.e. non-experts (public engagement, gender balance and diversity).
- Creating conditions for attracting and engaging young people and other underrepresented social groups in science and research (gender balance and diversity)
- Non-discrimination based on gender, ethnical origin or religion in all aspects and stages of the research (gender balance and diversity).
- Communicating all new knowledge through open-access/open-science practices (open access/open science).

ARC Fund's research practices and objectives have been therefore aligned with the RRI keys such as public engagement, open access, ethics and gender equality (included in the larger frame of non-discrimination and diversity) from the earliest years of organization's activity. Over the past six years, as a result of participation in several H2020-funded projects, RRI has become an increasingly important topic in the organization, necessitating a more formal uptake of the concept on all levels of management and among the research staff. Due to its small size, ARC Fund has no specific department, team or position dedicated to any of the RRI keys, nor the RRI framework as such. The ETHNA System project therefore provided an excellent opportunity for implementing a comprehensive ethics governance structure for conducting socially responsible and relevant research.

3 ETHNA System Implementation Process at ARC Fund

Based on the mapping process described briefly in the previous section, a detailed Action Plan for implementing the ETHNA System at ARC Fund has been developed. While the Action Plan roughly followed the ETHNA System Guide (González-Esteban et al. 2021) and the ETHNA Lab Methodology (Vedel Neuhaus et al. 2022), numerous modifications were made to adjust it to the needs and realities of a small research organization. The ETHNA Lab Methodology envisages six steps (Planning, Construction, Consultation, Refinement, Testing and Review). ARC Fund's Action Plan covered all six steps, and included a detailed overview of the necessary actions, along with the timeline, the list of concerned staff members, and progress and performance indicators for monitoring

and evaluating the process. During the implementation process it became clear that the Plan was too ambitious for such a small research organization, and certain modifications were made, mostly in the way of simplification of some of the planned activities or their rescheduling for the post-implementation period.

3.1 Selection of the ETHNA System Implementation Level

The ETHNA System has three level of institutional commitment (see González-Esteban et al. 2021).

- Level 1: The organization appoints the RRI Office or the RRI Officer and supports its activity. The RRI Office(r) will be in charge of disseminating the ETHNA System concepts, promoting awareness of principles and values, establishing activities and performance indicators for the three-year Action Plan for continuous improvement, and monitoring the progress of the ETHNA System in the organization through progress indicators.
- Level 2: The organization appoints the RRI Office or the RRI Officer and implements some of the Column Blocks (the Code of Ethics and Good Practices in R&I, the Ethics Committee on R&I, the Ethics Line). The Action Plan should incorporate at least one of the four major RRI keys: research integrity, gender perspective, open access, and public engagement.
- Level 3: The organization fully develops the ETHNA System. It designates the RRI Office(r), implements all three Column Blocks and applies a proactive attitude in all RRI key areas: research integrity, gender perspective, public engagement, and open access.

At the meeting with the senior management of the organization, it was decided that the second level of commitment was the most appropriate for ARC Fund. To kick-start the process, the Lab Manager and the members of the working group to write the Code of Ethics and Good Practices in R&I were appointed. It was decided that the Code should cover all four RRI keys, which are relevant for ARC Fund's work: research integrity, public engagement, gender equality and diversity, and open access. Finally, the management also approved the idea to set up the position of the RRI Officer and to establish a three-member Research Ethics Board. The Ethics Line was dismissed, as it was deemed non-applicable for a small organization.

3.2 The Lab Manager

In accordance with the guidelines provided by the ETHNA Lab Methodology (Vedel Neuhaus et al. 2022), the Lab Manager[6] was appointed. The Lab Manager is a person responsible for the planning, facilitating, coordinating and monitoring of the entire process. He or she also provides support to all other actors for performing their responsibilities and manages the communication with internal and external stakeholders.

The role of the Lab Manager is essential for the success of the implementation process. The experience of ARC Fund (but also confirmed by other cases presented in this volume) shows beyond doubt that the ETHNA System cannot be implemented if the entire process is not planned, executed and managed by a dedicated and committed person that is willing and able to take up this considerable responsibility.

Although this is not an obligatory precondition, such expert should preferably be well-versed in the RRI concept, or at least have considerable experience with some of the RRI keys, given that this policy framework plays a central role in the structure of the ETHNA System. The unfamiliarity with the topic might make it difficult to comprehend and follow some of the processes applied in the methodology. In the case of ARC Fund, an expert with considerable previous experience, obtained through participation in several RRI-focused research projects, was selected as the Lab Manager.

Good knowledge of the RRI framework is not the only skill that would ease the work of the Lab Manager. Other necessary competencies include good organizational, communication and leadership skills.

In addition, the Lab Manager should also be respected by the management, his or her colleagues and other internal and external stakeholders of the organization in order to successfully advocate for the implementation of the ETHNA System. It should not be taken for granted that the management, colleagues and other stakeholders will be familiar with the RRI concept, nor that they will be immensely eager to embrace the new policy and practice in the organization. This does not mean that a new ethics governance system would be rejected as something unnecessary or unwanted. Rather, this remark is a reminder that institutional change is a slow process that necessitates a continuous and resolute effort of a dedicated person (in the case of a small research organization) or a team (in the larger organization).

In short, the selection of an appropriate Lab Manager is a crucial step in the process. The candidate should have a certain moral authority within the organization, which means that junior staff members would be suitable for the task only in exceptional cases. At the same time, the Lab Manager should not be a member of the organization's management, because this would reinforce the possible perception among the research staff that another set of obligations is being imposed on them from the top. To be most successful, the Lab

[6] The term Lab Manager is used in this chapter for the purpose of consistency with the terminology applied in the ETHNA Lab Methodology. The process of implementing the ETHNA System at ARC Fund and other organizations presented in this book was referred to as ETHNA Lab. Hence, the person responsible for the process was named the Lab Manager. In the context of other organizations, which were not part of the ETHNA System project, and which would like to implement the ethics governance system described and proposed by this book, more appropriate terms for such a person would be Implementation Manager, Ethics Manager, Ethics Officer or RRI Officer.

Manager should therefore have a good communication with the management and enjoy their support, but should be in the position to act with considerable autonomy and not be influenced by the particular interests of the management. This might not be a huge problem in large organizations, but in small ones, such as ARC Fund, the list of potential candidates can be exceptionally short, making the selection and appointment of the Lab Manager a tough nut to crack.

Once this initial hurdle has been successfully dealt with, the newly appointed Lab Manager can begin the work. During the implementation process at ARC Fund, the Lab Manager was responsible for the following tasks:

- Drafting of the Action Plan;
- Presentation of the Action Plan to the senior management, along with the recommendation which level of commitment to the ETHNA System best matches the needs, objectives, and priorities of ARC Fund;
- Coordination of the Working Group, tasked with the writing of the Code of Ethics and Good Research Practices;
- Editing of all draft versions of the document;
- Planning and organization of all activities, which involved internal and externals stakeholders – interviews, focus group, meetings;
- Regular monitoring of the process and its evaluation.

That being said, it is essential to underline that the Lab Manager should not be expected to carry the entire process on their shoulders. The example of ARC Fund clearly shows that an active support from other relevant actors within the organization is obligatory for the successful development and uptake of the ethics governance structures in the organization. Insufficient involvement of other internal stakeholders may considerably slow down or even arrest the implementation process.

3.3 The Action Plan

The Action Plan has foreseen 21 different activities. Of these, 11 were fully completed, 4 were partially achieved, and 6 were not performed during the implementation process (but remain on the agenda for 2023).

Actions achieved within the planned timeframe:

- Meeting with senior management (organized in January 2022).
- Establishment of a Working Group to write the Code of Ethics and Good Practices in R&I (January 2022).
- Mapping of external stakeholders (February 2022).
- The first draft of the Code of Ethics and Good Practices in R&I (February 2022).
- Recruitment of external stakeholders (April 2022).
- Workshop with internal stakeholders (conducted in June 2022).
- The second draft of Code of Ethics and Good Practices in R&I (July 2022).
- Workshop with external stakeholders to present and promote the ETHNA System and the draft of the Code of Ethics and Good Practices in R&I (organized in July 2022).

- Evaluation of the ETHNA System implementation process (the first comprehensive evaluation of the process has been completed in November 2022; regular evaluation activities will continue as part of the annual evaluation and reporting process at ARC Fund).
- Evaluation report about the ETHNA System implementation (November 2022)
- Annual reporting on RRI performance in ARC Fund's Annual Report (December 2022 and will continue annually).

Actions partially completed (by the end of 2022):

- Final version of ARC Fund's Code of Ethics and Good Practices in R&I: The second and close-to-final draft was presented to the senior management in August, but was rejected with the explanation that it should be thoroughly revised and split up into four different documents: Code of Ethics and Good Research Practices; Gender Equality Plan; Open Access Policy; and Public Engagement Policy. By November 2022, only the Gender Equality Plan has been finalized and approved. The other three documents have not been finished before the end of the implementation process (November 2022). They are expected to be approved by the Board well before the completion of the ETHNA System project in June 2023.
- Designate RRI Officer: Only informal decision was taken in January 2022. The official nomination and formal approval by the Board of Trustees have not been made during the implementation. However, this is expected to occur in the spring of 2023.
- Establish Research Ethics Board: Similarly to the position of the RRI Officer, only an informal consent has been given by the senior management, but no actual steps were taken to select the members of the Research Ethics Board and make it operative. The expectation is that the procedure will be completed after the endorsement of the Code of Ethics and Good Research Practices.
- Set up the Advisory Group to discuss the second draft of the Code of Ethics and Good Practices in R&I: During the implementation, it was decided that due to the small size of ARC Fund, creation of an Advisory Group would not make sense and instead the entire staff was invited to read and comment on the draft document – this has been accomplished in the period May-June 2022.

Actions not performed by the end of 2022:

- Training on research ethics for all members of staff: It is planned to take place after the Code of Ethics and Good Research Practices is finalized and endorsed by the Board of Trustees (presumably by June 2023).
- Training of research staff on public engagement methods: It is planned to take place after the Public Engagement Policy is finalized and endorsed by the Board of Trustees (presumably by June 2023).
- Training of research staff on gender equality and diversity issues in research: It is planned to take place in April 2023.
- Training of research staff on open access issues: It is planned to take place after the Open Access Policy is finalized and endorsed by the Board of Trustees (presumably by June 2023).

- Review workshop with internal stakeholders: It was not conducted due to delays in the implementation process and because it was assessed that there were no new and important developments to review.
- 'RRI' dialogues with relevant external stakeholders: This event was not conducted because it was decided that external stakeholders cannot really contribute to the process in a productive way.

3.4 The RRI Officer

After the completion of the Living Lab process at ARC Fund, the Lab Manager was expected to be nominated for the position of the RRI Officer. However, as noted above, this intention has not been carried out by the time of writing of this chapter and remains to be enacted in the near future (most likely in late spring 2023).

The precondition for the appointment of the RRI Officer is the official endorsement of the revised Code of Ethics and Good Research Practices. The draft version of the document envisages for this position a staff member with sufficient knowledge of RRI concept and rich experience with RRI-related issues. The duration of the term shall be two years, with a possibility of re-election for another term (two consecutive terms in total).

Once appointed, the RRI Officer will act as the main contact point within the organization for all issues pertaining to the implementation and promotion of the Code of Ethics and Good Research Practices, including all activities related to research integrity, gender equality and diversity, open access and public engagement. The RRI Officer will be responsible for promoting the culture of research integrity at ARC Fund; communication and cooperation with the Research Ethics Board; planning and implementation of training and communication activities pertaining to the four RRI areas; and monitoring of RRI-related progress through an annual internal evaluation. The RRI Officer will also receive and review anonymous signals and complaints submitted by the members of staff, related to ethical, gender or diversity issues.

3.5 The Code of Ethics and Good Research Practices

A special Working Group was set up to write the Code of Ethics and Good Practices. Several drafts have been produced, but the approval and endorsement of the final document is still pending. To date, three drafts of the Code of Ethics and Good Practices in R&I have been written. The first two drafts were written as a comprehensive document, divided into four main parts, each covering one of the four RRI keys (research ethics, public engagement, gender equality and open access).

The third draft actually consists of four separate documents, each dedicated to one RRI key, according to the demand from the senior management. Of the four documents, only the Gender Equality Plan has been endorsed by the end of 2022. The other three documents are expected to be finalized and approved in the spring of 2023.

The Code acknowledged the organization's deep commitment to the RRI principles and built on the existing organizational documents such as the Code of Conduct, the Code of Ethics and the Child Protection Policy. Once officially approved, the Code is expected to further enhance ARC Fund's adherence to the RRI principles, setting up a flexible ethics governance system for the management of organization's research activities.

3.6 Scope and Form of Institutional Changes

At the time of writing of this report, no actual sustainable institutional changes have taken place yet. Due to the slower than envisaged progress with the implementation, the tangible institutional changes will only become noticeable in 2023. The senior management (the Board of Trustees) has confirmed its commitment to establish the positions of the RRI Officer and to set up the three-member Research Ethics Board, but this has not yet been formalized. An official decision about both bodies is expected by mid-2023 at the latest.

4 The Main Barriers and Drivers of the Implementation Process

Learning about the obstacles that had to be tackled and about the assets that supported the process can be of substantial assistance to organizations planning the application of the ETHNA System. In the first place, knowing about the potential barriers and drivers in the process can help to foresee the implementation risks and to plan how to overcome them. Secondly, it sheds light on the necessary conditions required to support and implement the new ethics governance system. Thirdly, this is a crucial step towards a realistic assessment of the viability of the ETHNA System.

Each implementing organization has encountered different barriers and drivers, although several common traits can be outlined (see Chapter 4: A Global Assessment of the Implementations: specificities, barriers and drivers of a complex process, in this volume). The barriers and drivers that characterized the implementation at ARC Fund are discussed below.

4.1 Barriers

The planning stage was smooth and no significant barriers have been observed. The main challenge that had to be tackled during the preparation of the Action Plan was how to ensure that the Plan was relevant for all three thematic programs of ARC Fund, as they have very different foci and work with different stakeholder groups.[7] After the consultation with the program leaders, it was established that RRI could provide a common frame for the three programs, as most keys were already implicitly and sometimes explicitly embedded into their activities. Once finalized, the Action Plan was approved by the management of the organization.

Additional challenge was the lack of mechanisms or structures that would support the coordination of ethical practices and considerations among the three programs of ARC Fund.

[7] ARC Fund has three thematic programs: (i) Science, Technology and Innovation Policy Program (STIPP), (ii) Safer Internet Centre Program (SIC), and (iii) Innovation and Business Support Program (IBS). STIPP is the program most committed to the principles of RRI. All activities of STIPP aim to strengthen the link between research, society, and policy, and mobilize different actors and societal stakeholders in the exchange of knowledge, experience and ideas. SIC works to enhance the digital literacy of children and adults through various training, engagement and other initiatives. IBS promotes innovation and entrepreneurship culture in Bulgaria and helps Bulgarian SMEs to establish themselves on the global market.

The actual implementation of the ambitious Action Plan was much slower than expected. A combination of factors has undermined its timely execution. To start with, ARC Fund has no departments or bodies that would be responsible for the governance, management, and evaluation of policies and practices, pertaining to different RRI keys or the entire framework. This means that all people involved in the implementation had to accommodate these tasks within their very busy work schedule. In general, the considerable workload and the time constraints related to the project-based work, which is bounded by strict deadlines, were the main barrier, which affected the timely implementation of actions foreseen in the Plan.

The heavy workload of ARC Fund's research staff also made it difficult to conduct the planned workshops with the internal stakeholders. The Plan was overly ambitious and envisaged three workshops and four trainings with participation of the organization's employees.[8] Looking back at the process, it was completely unrealistic to expect that seven events can be indeed organized in a time span of nine months in an organization with less than 15 employees. Additional challenge was the COVID situation, as quite a few researchers worked from home, making it almost impossible to agree upon the appropriate dates and format for the events. Instead of the workshops, eight interviews were conducted and one focus group with nine participants was organized. The trainings, however, remain on the agenda and will be organized after the Code and other documents are finalized.

Additional barrier was a certain overexposure to the RRI topic among the staff. As already mentioned, ARC Fund has been involved in several RRI-focused projects since 2015. Due to the small size of personnel, practically all staff members have been involved in some capacity in RRI activities, either as researchers or as participants in interviews, workshops or focus groups. This created a noticeable discontent among some researchers for being asked to discuss "the same issues" all over again.

The organization has no practice of monitoring and evaluating the performance of ARC Fund's staff members in terms of compliance with the research ethics standards. Performance on research ethics has also not been reported in ARC Fund's Annual Report on activities. This represented a modest barrier inasmuch as it is never easy to change the established working practices or introduce new ones in the organization. This is expected to change in the future, as regular reporting on RRI progress and performance has been included in ARC Fund's Annual Report.

4.2 Drivers

The planning and preparation of the Action Plan were eased considerably by the knowledge and experience gained by participation in H2020-funded projects focused on RRI. Especially important driver was the work on the RRI-Practice project,[9] during which the

[8] The initial workshop for obtaining opinions and recommendations for content and form of the Code; mid-process workshop to present the draft Code; final review workshop to collect recommendations regarding the future status of ETHNA System at ARC Fund; training on research ethics; training on public engagement methods; training on gender equality and diversity issues in research; training on open access issues.

[9] See https://cordis.europa.eu/project/id/709637.

idea emerged that it would be highly beneficial for ARC Fund to formalize its commitment to the RRI principles in a comprehensive document. The suggestions to establish the Ethics Board and to regularly monitor and evaluate organization's performance in this area also came up during the involvement in this project. However, due to the workload and commitments of relevant internal stakeholders, these plans never came to fruition. The ETHNA System project provided an excellent opportunity to upgrade and finally implement the original idea.

The implementation process was beyond doubt facilitated by the fact that some RRI keys have already been included in the existing ARC Fund's documents, which regulate the work at the organization (the Statute, the Code of Conduct and the Code of Ethics). These documents served as the basis and inspiration for the writing of the Code of Ethics and Good Practices in R&I, along with the ETHNA Office Toolbox and the ETHNA System Guide to the Ethical Governance of RRI.

Additional driver that highlighted the necessity for the ETHNA System implementation were the requirements of the funding programs – for example the Open Access requirement for the EC funded projects, and the obligation for organizations wanting to participate in Horizon Europe projects to have the Gender Equality Plan.

The small size of the ARC Fund collective means that everyone knows all their colleagues personally, which helped with the communication and organization of the consultation activities. The staff members who participated in the interviews and focus group were also eager to take the opportunity to make constructive suggestions and recommendations – something they are not able to do very often due to the daily rush and tasks everyone at ARC Fund is affected by.

The input and comments from the staff and from several external stakeholders were an important driver for the editing and refining of the first two drafts of the Code of Ethics and Good Practices in R&I.

5 Good Practices

The continuity and consistency in engagement with the RRI framework has immensely eased the work on the ETHNA System implementation at ARC Fund. The accumulated experience and knowledge on RRI have made it easier to plan and prepare the process, and to identify appropriate progress and performance indicators for its evaluation. All people involved in the process had good knowledge and substantial experience with RRI related themes and there was no need to convince them about the benefits of adopting the ethics governance structure in the organization.[10] The work was undertaken in the spirit of mutual cooperation and support.

The interviews and the focus group clearly demonstrated how important it was to provide a forum for the employees to share their concerns and suggest solutions. Most of the interviewees welcomed the development of the Code of Ethics and Good Practices in R&I, but noted that efficient integration of RRI into the existing processes within the

[10] The previous knowledge and experience with RRI should be seen only as a facilitating factor and not at all as the necessary condition for small organizations that want to adopt ETHNA System. Lack of prior experience and knowledge of RRI is not a barrier, it only means that in most cases, more effort will be needed to explain and popularize the concept.

organization will be a slow process, and the drafting of the Code can only represent the first step.

6 Conclusions and Recommendations

6.1 Conclusions

- The ETHNA System implementation process at ARC Fund has validated the Toolbox to Implement ETHNA System and the ETHNA System Guide to the Ethical Governance of RRI as practical documents that provide useful instructions for implementing an ethics governance system in a research performing organization.
- In contrast, the ETHNA Lab methodology is only conditionally applicable in small institutions such as ARC Fund. There is only so much you can do with a pool of about 15 people, especially in the COVID period, when many staff members have been working from home. The ambitious cycle of several workshops and trainings foreseen in the methodology demands having at your disposal a large number of relevant internal stakeholders, otherwise the same people are asked time and again to contribute to or participate in a sequence of events. This inevitably leads to fatigue and impression that additional bureaucratic burden is being imposed upon the researchers.
- The process of implementation at ARC Fund was off to a promising start, bolstered by the fact that the core team involved in the process has been well-versed and experienced in the RRI matters and in the conduct of different necessary activities (interviews, focus groups, workshops, document writing). The declared support of the senior management also gave initial boost to the process.
- However, as time was passing, the implementation process started to stumble. The core team, being overwhelmed by a heavy workload on multiple projects, began to lose steam. The originally planned deadlines were postponed, and some of the crucial elements of the new ethics governance system have not been put in place before the end of the process in November 2022. They nevertheless remain on the organization's agenda and are planned for the first half of 2023.
- In hindsight, it is quite obvious that the original Action Plan has been overly ambitions – in number and type of activities and indicators. At the same time, the foreseen timeframe was too short. Nevertheless, it remains feasible that most of the objectives from the Action Plan will eventually be achieved – only several months later than planned.

6.2 Recommendations

- The importance of the Lab Manager for the smooth and timely implementation of the process can hardly be overstated. In organizations such as ARC Fund, which are 100% project funded and have no financial or other resources to support a position such as the RRI Officer, the ETHNA System can be introduced only if a dedicated and committed person is willing to take up the responsibility to plan, execute and monitor the process. Such expert should preferably have considerable experience in the RRI topics (or at least some of the RRI key areas), given the central place of the RRI principles in the structure of the ETHNA System.

- The Lab Manager should be respected by management, colleagues and other internal and external stakeholders of the organization. Organizational, communication and leadership skills are essential for the proper management of the working groups, involved in preparation of different documents, for "selling" the ETHNA System to the management, and for engaging the staff in different activities (interviews, focus groups and trainings). Without such an RRI champion, the process will be stillborn.
- Of course, the Lab Manager should not be expected to do everything alone. A genuine commitment and support from the management and the researchers is crucial. The process must not be directed top-down, because it is likely to be perceived by the lower ranks of personnel as additional unwanted and unnecessary burden. Instead, the researchers need to be convinced about the benefits of adhering to the ethical principles of conducting science. The management of the organization must find a proper way to stimulate the personnel and engage them in the co-creation process for reforming the "business-as-usual" approach to research in the organization.
- The successful uptake of the ethics governance system is only the first step. It is not enough to make the researchers aware of the new rules and requirements through presentations and trainings. An adequate monitoring and evaluation system is necessary to ensure that the everyday procedures and activities are aligned with them, otherwise the rules might remain only letters on a paper.
- Another issue that needs to be considered are the possible unexpected outcomes that might arise after the implementation of new structures such as the ETHNA System. The new documents such as the Code of Ethics or the Gender Equality Plan can provoke debates or considerations nobody in the organization has been thinking about earlier – for example, the gender issues, work related correspondence outside the business hours, or the erasure of sensitive data after the research has been completed. The responsible bodies, such as the RRI Officer and the Research Ethics Board need to be flexible and responsive enough to address these questions whenever they arise.
- Finally, it can be noted that external incentives, such as the EU funding requirements (e.g. in case of the Gender Equality Plan) or participation in a project such as ETHNA System, can kickstart the process and even assure that certain objectives are implemented. However, this is not enough to guarantee the sustainability of the institutional changes. In the beginning, the ETHNA System means above all more obligations for the staff, as they need to learn and abide by the new rules. The ETHNA System can only become a supportive and empowering instrument if the organization manages to align it with the context and needs of the employees. Overcoming the initial distrust, reluctance or even opposition is essential for engaging the staff in the process and triggering a stable institutional culture, which has internalized the ethics governance system. The sustainability of the process depends on the successful transformation of a top-down initiative into a genuine co-creation process "owned" by all stakeholders of the organization.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission.

References

González-Esteban, E., Feenstra, R., Calvo, P., Sanahuja Sanahuja, R., Fernández Beltrán, F., García-Marzá, D.: The ETHNA system - a guide to the ethical governance of RRI in innovation and research in research performing organizations and research funding organizations. Deliverable 4.2 Draft concept of the ETHNA System. ETHNA Project [872360] Horizon 2020 (2021). https://ethnasystem.eu/wp-content/uploads/2022/07/D4.2_ETHNA_2022_guide_220210_incl_toolbox_neu.pdf

Vedel Neuhaus, S., et al.: ETHNA System Implementation Co-design Requirements Guiding Paper – The ETHNA Lab. Deliverable D5.3 ETHNA Project [872360] Horizon 2020 (2022). https://ethnasystem.eu/wp-content/uploads/2022/05/D5.3_ETHNA_lab-method-guide.pdf

Implementing RRI in a Research and Innovation Ecosystem

Luis M. Camarinha-Matos(✉), Filipa Ferrada, and Ana Inês Oliveira

NOVA School of Science and Technology, UNINOVA - CTS and LASI, NOVA University of Lisbon, Campus de Caparica, 2829-516 Monte de Caparica, Portugal
{cam,faf,aio}@uninova.pt

Abstract. New organizational forms are emerging today at all levels of society, and more and more research is conducted in dynamic collaborative networks or ecosystems. Unlike traditional research centers, these new types of organization are very dynamic, with fluid boundaries, and volatile in terms of membership. This characteristic requires that more attention be paid to research ethics and RRI. This work reports on an implementation process carried out in a research and innovation ecosystem according to the principles and guidelines proposed by the ETHNA project. The process, its barriers and drivers are described, and finally, learned lessons and recommendations are presented.

Keywords: Responsible Research and Innovation · Research Ethics · Research and Innovation Ecosystem · Living Lab · Ethical governance

1 Introduction

In recent years, there is a growing concern about the multiple dimensions of research ethics and responsible research and innovation (RRI) [1, 2] among various stakeholders such as funding agencies, publishers, research centers, scientific societies, and the research community at large. However, the level of awareness of the available principles and mechanisms is not homogeneous across all fields. For instance, while these issues have been at the center of attention of communities involved in health-related or biomedical research for a long time, the situation in engineering research is somewhat different and has only more recently begun to be discussed. The exception is perhaps the case of the technology management, environment engineering, and innovation subfields which have included the topic of RRI in their agendas for some time now [3]. The fact that a substantial amount of the RRI literature produced by social and health-related scientists is seen by engineering researchers as too theoretical and even using a hermetic language also contributes to this difference. Nevertheless, even in engineering and technology development fields the situation is changing, at least as a result of the pressure from research funding agencies [1, 4] and publishers. Often research programs require a link to the sustainability dimension, e.g., by addressing the UN Agenda 2030 for sustainable development [5]. For instance, the various sub-items of Goal 9 of this

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 87–108, 2023.
https://doi.org/10.1007/978-3-031-33177-0_6

agenda, "*Build resilient infrastructure, promote inclusive and sustainable industrialisation and foster innovation*" clearly relate to a perspective of RRI in manufacturing. The growing maturity of Artificial Intelligence and its fast adoption in novel technological developments is also motivating extensive discussion on ethics and technology [6]. Furthermore, the growing societal demand for accountability, particularly when research is supported by public funds, and the impact of widespread Internet/social networks reporting cases of plagiarism, fabrication, and other forms of research misconduct, including the emergence of "criminal science publishing gangs", are acting as a wake-up call to the entire research community [3, 7]. The growing importance of "data science" also raises the need to understand the responsibility of data management and develop proper data governance mechanisms [8].

On the other hand, in recent decades, new forms of organizing research have emerged, notably leveraging different forms of collaborative networks [9–11]. Compared to traditional organizations with very precise "boundaries", such as universities, research institutes and research centers, the more fluid, and dynamic networked organizations such as research and innovation ecosystems [12, 13] bring a new level of complexity that, at the same time, makes the establishment of appropriate RRI principles and mechanisms more crucial for a healthy collaboration and sustainability of those organizational forms.

This chapter focuses on the implementation of research ethics and RRI dimensions in a collaborative research and innovation ecosystem in Portugal. It thus discusses the experience gained from adopting the ETHNA system and guidelines. The lessons learned are likely to benefit similar research and innovation ecosystems and research networks in general.

2 Research and Innovation Ecosystem Structure

When trying to increase the acceptance of RRI it is important to understand and take into account the organizational context and how to "manoeuvre inside the organization" in order to succeed and overcome resistance [14]. The RRI implementation case here reported was carried out during 2022 at the Center of Technology and Systems (CTS) of UNINOVA (https://cts.uninova.pt/), hereafter referred to as UNINOVA-CTS, which is a research center recognized by the Portuguese government agency for science and technology (FCT – "Fundação para a Ciência e Tecnologia").

2.1 Ecosystem Structure

UNINOVA-CTS, like many other research centers in Portugal, is a kind of research and innovation ecosystem, including researchers who have an employment contract with one of the following entities: NOVA School of Science and Technology (a faculty of the NOVA University of Lisbon, to which most members are connected), ISEL/Polytechnic Institute of Lisbon, Polytechnic Institute of Setubal, and Polytechnic Institute of Beja. Most of these members are part of the academic staff of the mentioned institutions, where they teach, but carry out their research activities in the context of UNINOVA-CTS as a result of a cooperation agreement. This includes a total of about 75 researchers with a

PhD. In addition, this ecosystem also includes about 80–100 PhD students who carry out their research work for their theses at this center. From an administrative point of view, CTS is hosted and managed by UNINOVA, a legal, not-for-profit entity, also part of the "periphery" of NOVA University of Lisbon (Fig. 1).

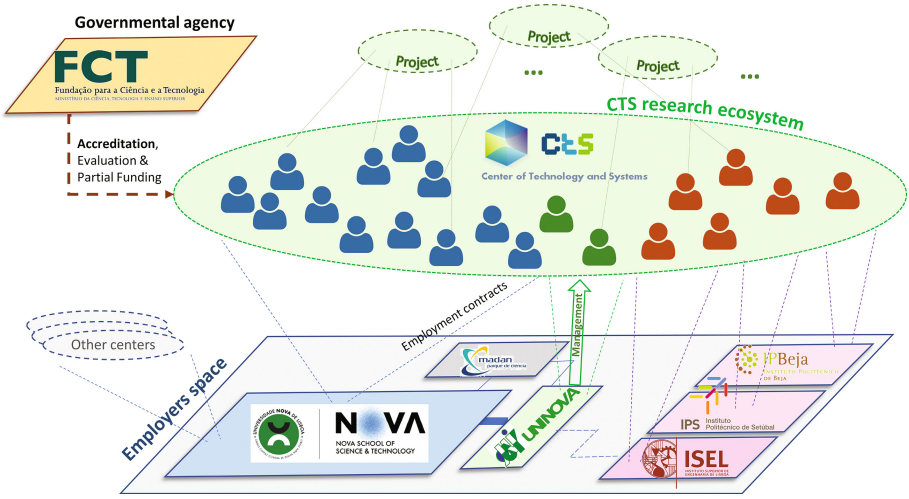


Fig. 1. The CTS Research and Innovation Ecosystem

Research at UNINOVA-CTS is organized in projects, which are typically funded by international (e.g., European Commission and European Space Agency) and national programs. These projects are typically carried out in consortia involving academic and industrial partners. Some researchers may be involved in the creation of spinoffs to exploit research results. These spinoffs are typically hosted, at a preliminary phase, in the Madan Park incubator, also located in the vicinity of our campus.

UNINOVA-CTS research addresses engineering systems with a cyber-physical dimension in the broad area of information and communication technologies, including modelling and design, development of support technologies and methods, proposition of adequate governance models, application, and assessment. The center covers a wide spectrum of knowledge areas in electrical and computer engineering and aims to further knowledge and technology development towards cognitive and collaborative cyber-physical systems, while pursuing interdisciplinary integration (Fig. 2).

The center is concerned with contributing to contemporary societal challenges, including a strong component of applied research in industry and services, guiding its action through the continuous search for excellence in research and effective value creation and valorization of research results. This aim also includes a strong commitment to the training of young researchers and early career researchers, and to having an active presence in international networks, contributing to strategic research agendas, and engaging with societal stakeholders. However, despite the mentioned objectives, the level of awareness and implementation of RRI at the begin of this initiative was relatively low.

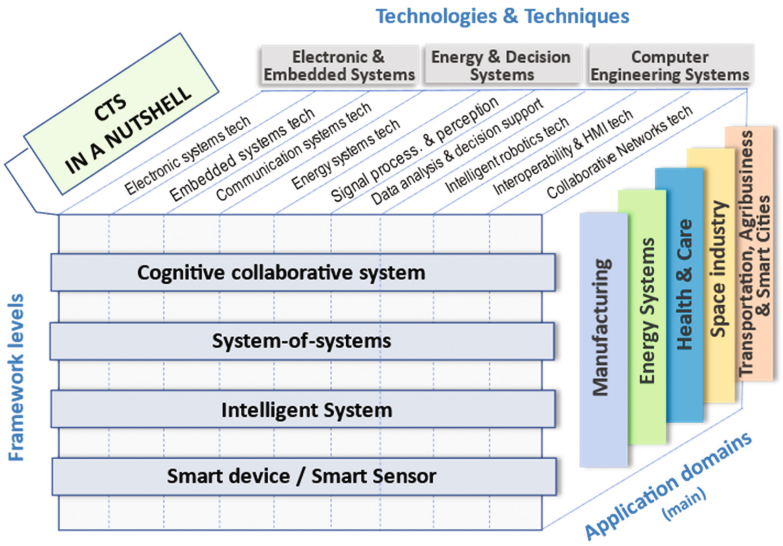


Fig. 2. UNINOVA-CTS scope

According to its mission, UNINOVA-CTS is committed to: (i) Planning and conducting high quality research on advanced engineering systems [excellence in research], (ii) Creating value and societal impact with research results [excellence in society engagement], (iii) training the future generation of researchers in the field [excellence in education]. The center pursues its objectives guided by the following values: (a) academic honesty and responsibility, (b) appreciation of excellence, (c) appraisal of creativity and entrepreneurial spirit, (d) respect for individual intellectual freedom, and (f) attention to societal concerns. Although the presence of ethics governance mechanisms was low at the beginning of this implementation, such announced values indicate a pre-disposition to implement RRI.

2.2 Complementarity Approach

Each researcher at UNINOVA-CTS is indeed subject to various "RRI spaces" (Fig. 3), namely:

A. *Employer's RRI space*: first, the researcher must comply with the code of ethics and other RRI principles of the employer which, nevertheless, are not always extensively disseminated.
B. *CTS RRI space*: then he/she needs to comply with the RRI principles of the CTS research ecosystem.
C. *Projects' RRI space*: each time a researcher is involved in a project, he/she needs to comply, during the duration of the project, with the RRI principles defined by the funding agency for that specific project/program.
D. *Scientific society code of ethics*: finally, most researchers are members of international and national scientific and technical societies (e.g., Institute of Electrical and

Electronic Engineers (IEEE), International Federation of Information Processing (IFIP), International Federation of Automatic Control (IFAC), Society of Collaborative Networks (Socolnet), National Engineers Association) and as such they need to comply with the code of ethics of these societies.

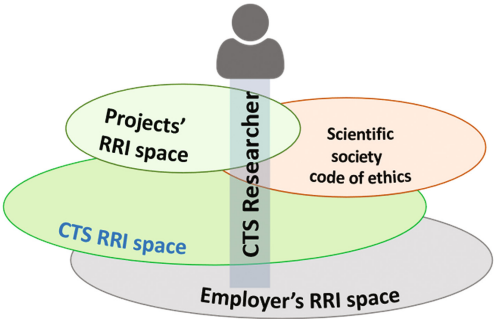


Fig. 3. Coping with multiple RRI spaces

More specifically, regarding each "space":

A) All mentioned (academic) employers are public institutions and as such follow general rules of public entities, including gender equality, conflicts of interest, rules against plagiarism, etc., but such principles should deserve more structured dissemination. In terms of organizational units: NOVA University of Lisbon has an *Ethics Council* that acts as an advisory body to the rector, but so far with limited interaction with the research community. Some seminars on *responsible data management* have also been organized for the directors of research centers. ISEL/Polytechnic Institute of Lisbon has no specific RRI committee or rules; at Polytechnic Institute of Setubal due to the existence of a health school, there is an ethics committee dedicated to health, but no general RRI principles for the entire institute; Polytechnic Institute of Beja has an *Ethics Committee* that acts as an advisory body with the mission to promote high standards of integrity, honesty, and best practices but most researchers are not well acquainted with it.

B) In its agreement with the accreditation agency (FCT – Foundation for Science and Technology), UNINOVA-CTS has stated that the center is committed to carry out its activities under the widely accepted principles of Research Ethics and Responsible Research. To prevent misconduct and bad practices, the "**European Charter for Researchers**" [15] was identified as a useful guide, namely along the principles and recommendations concerning: Research Freedom, Ethical Principles, Professional Responsibility, Professional Attitude, Contractual and Legal Obligations, Accountability, Good Practices in Research, Dissemination and Exploitation of Results, Public Engagement, Relationship with Supervisors, Supervision and Managerial Duties, Continuing Professional Development.

These principles are conveyed to PhD students and early-stage researchers through training actions. The IEEE authorship principles [16] have been promoted and frequently reminded to all researchers. At the start of this project, there were no specific organizational units devoted to RRI, being these issues dealt with at the management board level.

C) Some research funding programs impose general and specific ethical principles, such as the case of European Commission's programs in Horizon 2020 and Horizon Europe [17, 18], and national funding programs.
D) The mentioned societies usually have a code of ethics. One example: IEEE Code of Ethics, but in practice most researchers are not well acquainted with it.

Based on this context, Table 1 summarizes what we could identify as existing resources per RRI key.

Table 1. Pre-existing RRI resources for UNINOVA-CTS

RRI key	**RRI Space**	**Existing resources before starting the implementation process**
Research integrity	*Employers*	Ethics committee, but not widely known; DPO-GDPR sub-contracted
	Uninova-CTS	Some promotion of good research practices but lacking a formalized model and wider adoption; Ethical Code of IEEE regarding authorship is amply disseminated among researchers; Some promotion of RRI among PhD students
	Projects	Follow the general principles of Research Ethics; RRI is explicitly enforced by the funding organizations
	Scientific societies	Code of ethics of IEEE; Code of ethics of IFIP; Code of ethics of national engineers' association, but not widely disseminated
Gender perspective	*Employers*	General rules/laws of Portuguese Government
	Uninova-CTS	The center depends on the rules followed by the institutions that employ CTS researchers. Current gender distribution is not balanced but this is not the result of any discrimination, rather a reflection of the gender (un)balance in this scientific field
	Projects	Compliance with rules of funding programs
	Scientific societies	Some initiatives to encourage women in engineering (e.g., IEEE)

(*continued*)

Table 1. (*continued*)

RRI key	RRI Space	Existing resources before starting the implementation process
Public engagement	*Employers*	Office for Innovation Research and Impact Strategy (IRIS), a recent initiative at NOVA; Vice-dean for S&T dissemination
	Uninova-CTS	Involved in technology transfer and results exploitation actions; A good number of spin-offs (average of 1 per year in the last decade) originated from CTS members; Many dissemination events (ad-hoc) but not uniform in all areas of activity of the center
	Projects	Many projects require involvement of end-users
	Scientific societies	Use of existing channels in scientific societies to reach a wider engagement with society
Open access	*Employers*	Some repositories: PURE repository; IPL repository + open access journal
	Uninova-CTS	Promotion of knowledge sharing and publication in open access channels (but constrained by the high and fast increasing financial costs of open access publications)
	Projects	Open access required by funding organizations e.g., EC, FCT
	Scientific societies	Open access journals (e.g., IEEE)

These resources, which are disperse and of which the researchers are not fully aware were essential in determining the position of CTS with respect to RRI implementation needs and were used as the basis for reaching the level of commitment regarding the implementation of the ETHNA system [19] through the identification of the main goals and priorities.

It should also be noted that participation of researchers in UNINOVA-CTS is voluntary. The employment institutions encourage or demand that staff get involved in a research center, but researchers can choose the center/ecosystem they join and can move from one center to another. As a result, there is some volatility of membership regarding PhD holders. Regarding PhD (and MSc) students, they remain associated with the center for the duration of their studies and naturally leave after their graduation. Since the population of PhD students represents a little over half of the total membership of UNINOVA-CTS, this strongly contributes to a high volatility of the research population, requiring continued attention to the endogenization of RRI principles and practices.

2.3 Needs, Challenges and Opportunities

Considering the described characterization of UNINOVA-CTS, the following main needs were identified for research ethics and RRI implementation:

- Development of better structured information repositories/navigation map on RRI principles and tools.
- Adoption/establishment of straightforward implementation guidelines.
- Organization of a repository of templates/models/exemplary cases with particular focus on:
  - Ethical principles of authorship;
  - Conflicts of interest, namely in the context of creation of spinoffs and exploitation of results in interaction with industry;
  - Research data management;
  - Relationships between supervisors and PhD students;
  - Clear identification of added **value** of RRI for engineering researchers (of utmost importance);
  - Very flexible governance approaches. In a research ecosystem like UNINOVA-CTS (which is a kind of collaborative network), the creation of dedicated "organizational structures" for RRI would be an extra overhead, extremely difficult to implement. As such, the governance models proposed by ETHNA sound a bit "too bureaucratic" and require funds that are not available.

Since the organizational structure of UNINOVA-CTS (a collaborative and distributed research ecosystem) is common to many centers in Portugal, results from ETHNA with the above characteristics could be replicable to those centers.

As the Director of the center was engaged, from the beginning, in this implementation process, the departing situation points to a strong leadership. In what concerns the base, RRI norms and practices have not been effectively implemented at the institutional level. They are fragmented by each RRI space. As such, CTS-UNINOVA fits into the "strong leadership/weak base" quadrant of the ETHNA classification (Fig. 4).

Furthermore:

- The UNINOVA-CTS research center, due to its mission to promote excellent research and innovation practices and the commitments assumed with the Portuguese research funding agency, has already some awareness regarding RRI in all its key areas. However, making RRI awareness widespread and materialized in concrete rules and mechanisms used by all researchers still need a considerable effort.
- In this line, some initiatives have already started before ETHNA implementation, namely regarding research ethics and integrity, such as some promotion of research good practices or RRI awareness among PhD students. However, it is expected a long way to go because there is a lack of formalized models matching our context and the RRI norms and practices are not yet fully adopted by most of the CTS research members.

Fig. 4. ETHNA RRI institutionalization Quadrants and UNINOVA-CTS

– It is also noticeable that although there is good awareness of the RRI importance at the management board level, CTS researchers are quite busy with their own research and with acquiring new funded projects, and thus not motivated to play a proactive role or spend time with RRI implementation.

Based on the identified situation, it was decided to pursue a Level 2 implementation of the ETHNA system [19] (see Fig. 5). However, regarding pillar 2 (Ethics Committee & Ethics Line), due to the lack of specific funds, it was decided to follow a minimalist approach, just creating a RRI task force.

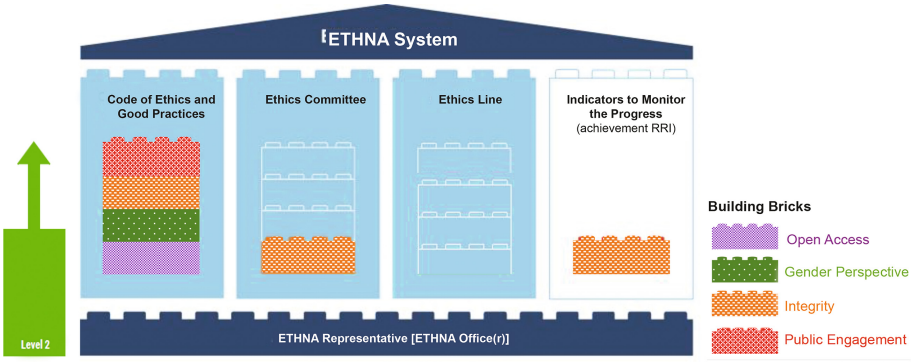


Fig. 5. UNINOVA-CTS Level of implementation commitment

3 The Implementation Process

The implementation of the ETHNA System is an iterative process (living lab) that consists of three main phases decomposed into six consecutive sub-phases as depicted in Fig. 6.

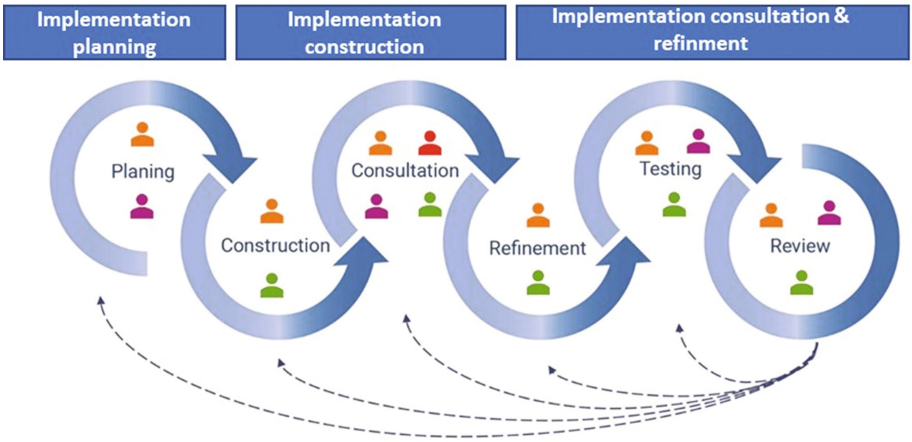


Fig. 6. ETHNA System implementation process

3.1 Phases of the Process

Although taking these ETHNA System guidelines as a general approach, it was necessary to make some simplification/adaptation of the process. This simplification was needed mainly motivated by the following reasons:

– Researchers are quite busy with their own work and they are not "RRI researchers" (i.e., they do not do research on RRI itself). Thus, it is not realistic to expect the involvement of researchers in too many meetings/workshops for RRI implementation.
– The process started during COVID-19 confinement period, which also limited interactions among researchers and other stakeholders.

Therefore, the main (adapted) steps were the following:

Implementation Planning. The main objective of this phase is to prepare and initiate the implementation process. The ETHNA System guide was slightly modified and adapted to the UNINOVA-CTS case, resulting in 4 main steps as depicted in Fig. 7:

- *Step 1: Contextualization.* Since members of UNINOVA-CTS are extremely focused on their own research activities it was necessary to find the right "political approach" for engaging people and adapting the ETHNA's "language" to a more contextualized and general discourse to motivate participants.

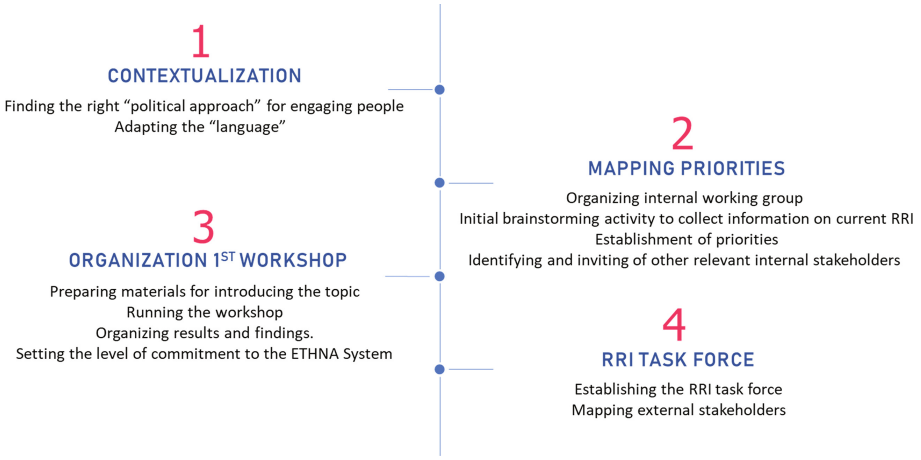


Fig. 7. Adapted phases of the implementation steps

- *Step 2: Mapping priorities.* The initial information on current RRI status at CTS was collected with the help of an internal working group through brainstorming. It was useful on one hand to understand the CTS-UNINOVA situation regarding the different RRI keys, and on the other hand to establish the goals and corresponding implementation priorities considering the available resources, the capabilities, and identified objectives. In this step it became evident that one of the first priorities was the organization of the RRI knowledge repository and the creation of a RRI taskforce. During this phase, other relevant internal stakeholders were also identified and invited to join the implementation activities.
- *Step 3: Organization of the 1st internal workshop.* This phase comprised the preparation of the materials for introducing the ETHNA System to the CTS members, on running the workshop and organizing the results and findings. As a result of this workshop, some adjustments were made to the priorities established by the internal working group. For instance, besides the issues of gender perspective, it was decided to consider a more general scope of inclusion. This was motivated by the diversity of nationalities and cultural/religious backgrounds of CTS members. It was also defined the CTS's level of commitment to the ETHNA System as illustrated by the blocks highlighted in Fig. 5.
- *Step 4: Establishing RRI task force.* Consisted of the identification and establishment of the RRI task force that is composed of 5 CTS senior researchers including the director of the center and one member of the board. At this stage a mapping of the external stakeholders was also elaborated.

Implementation Construction. This phase consisted of a comprehensive working activity on the implementation of the Code of Ethics and Good Practices in R&I for ETHNA's building blocks (Research Integrity, Open Access, Public Engagement and Gender Perspective). This activity was mainly conducted by the RRI task force and complemented with a second workshop with experts and internal stakeholders.

The recruitment of the external stakeholders for the next Consultation phase was also performed.

Implementation Consultation and Refinement. This step consisted in a workshop with internal and external stakeholders from research and education, business and industry, and civil society. The main objective of this workshop was to present the implementation status of Code of Ethics and Good Practices in R&I, create a discussion session and collect inputs from the different perspectives. After this consultation workshop a refinement phase was initiated aiming at refining the developed components of the CTS ETHNA System until a collective agreement on the content of those key documents was reached. For this stage, online feedback collection modality was used. The agreement was reached on the assumption that these are "live documents" that need to go through continuous improvement. For instance, the guidelines on open access are likely to evolve as the "open access market/channels" also evolves.

3.2 Main Outcomes

Considering the specific nature of the research center, in which most researchers are employees of other institutions (Universities and Polytechnic Institutes) and thus already subject to different ethical systems, the implementation of ETHNA system focused mainly on complementarities. As such, a number of key documents on various RRI aspects (Code of Ethics and Good Practices in Research and Innovation, Open Access Guidelines, Gender and Inclusion Equality Plan), illustrated in Figs. 8, 9 and 10, were elaborated. The included items and the level of detail in each document resulted from the identified priorities and considering the mentioned perspective of "complementarity".

A particularly useful instrument is the Open Access Guidelines. On one hand, all researchers feel the pressure from funding agencies to publish in open access (often a mandatory requirement) and to pursue principles of open science. But on the other hand, they have to face contradictory challenges, namely:

- The cost of open access publications, which puts an added pressure on researchers to acquire further resources.
- The fact that open access publications are still less prestigious than publications in "traditional" channels. The proliferation of predatory open access journals does not help. As a result, a considerable number of researchers do not value publications for which the authors have to pay to publish (the prevalent model behind open access) and this can even have a negative impact when researchers are evaluated. Having some practical guidance on how to deal with such "confusing situation" is perceived by all CTS researchers as a critical need.

Another outcome was the creation of awareness among the UNINOVA-CTS community by creating a specific section on RRI in the organization's web site and organizing some internal dissemination seminars.

Additionally, considering that all members of the center are mostly focused on their own research activities and no resources are available to create additional organizational

Fig. 8. Scope of the Code of Ethics and Good Practices document

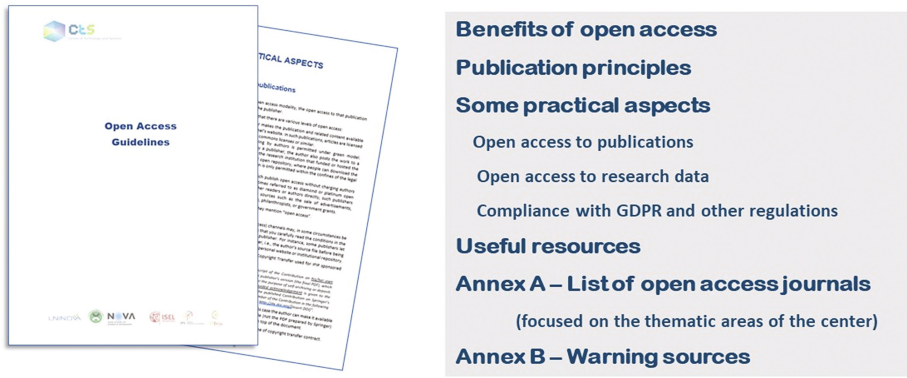


Fig. 9. Scope of the Open Access Guidelines document

structures, the organizational changes were limited to launching an RRI Task Force instead of a very formal RRI committee.

Finally, another result of this implementation has been the strengthening of training sessions for young researchers/PhD students. All our PhD students have an initial course on Scientific Research Methodologies and Techniques, which includes a specific unit on Research Ethics (Fig. 11, Unit 10). Beyond this specific unit, RRI principles and practices are emphasized in all other units. For example, when discussing Publication of Results (Unit 4), students are amply warned of unacceptable practices and behaviors in publishing and motivated for the benefits of conducting research according to proper standards. Even in the Introduction unit, RRI is present e.g., when discussing the relationship between the supervisor and the student.

Furthermore, we usually include some invited talks on research ethics and RRI in the annual DoCEIS conference that is organized for doctoral students [20].

Progress and performance indicators are included in the Table 2. A more extensive list of indicators can be found in [21].

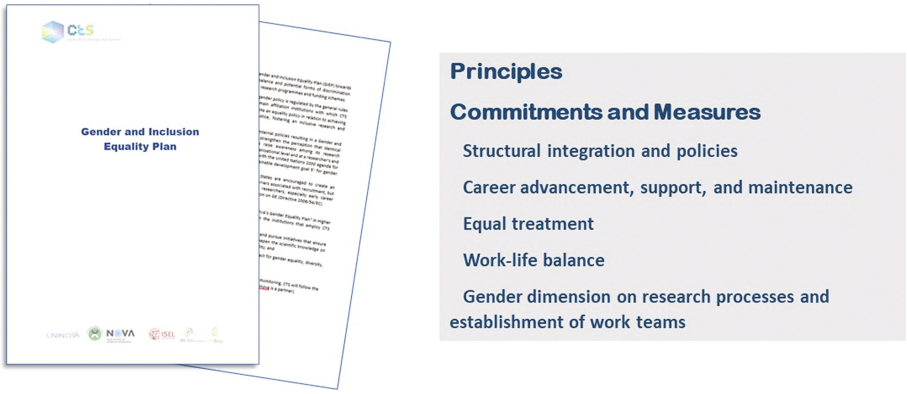


Fig. 10. Scope of the Gender and Inclusion Equality Plan document

Unit 1: INTRODUCTION
Objectives for a PhD, base concepts, types of research, relationship with supervisor
Unit 2: SCIENTIFIC METHOD
Overview of research methods, steps of the scientific method, engineering research and design science research method
Unit 3: LITERATURE REVISION
Information sources, information search, special sources, synthesis and critics
Unit 4: PUBLICATION OF RESULTS
Writing scientific papers, publication channels, evaluation procedures, citations
Unit 5: THESIS ORGANIZATION AND VALIDATION
Structure, research question, thesis contribution, validation of results
Unit 6: RESEARCH IN COLLABORATION
Types of projects and partnerships, requirements, collaboration spirit and constraints
Unit 7: PROJECT PROPOSAL PREPARATION
General structure of a proposal, typical example
Unit 8: RESEARCH PROJECT MANAGEMENT
Management structure, management principles, tools, risks, reporting
Unit 9: ASSESSMENT OF RESEARCH RESULTS
Phases of research and outcomes, research performance indicators
Unit 10: RESEARCH ETHICS
Ethical issues and behavior, responsible conduct, scientific practices and violation
Unit 11: INTELLECTUAL PROPERTY RIGHTS
Concepts, types, protection mechanisms, rights identification, rights transfer
Unit 12: ROADMAPPING AND FUTURE PLANNING (1)
Future planning objectives and approaches, concept of roadmapping
Unit 13: ROADMAPPING AND FUTURE PLANNING (2)
Roadmapping methodology
Unit 14: PROJECT PROPOSAL PREPARATION - EXAMPLES
Examples in different programs – H2020, ESA
Unit 15: PANEL – PROPOSALS NEGOTIATION EXERCISE

Fig. 11. Scientific Research Methodologies and Techniques course

Table 2. List of progress and performance indicators for the UNINOVA-CTS implementation

Progress Indicators	Achievement
Creation of RRI working group to elaborate CEGP	Fully
Decision of the coverage of CEGP	Fully
Launch a participatory process with stakeholders to discuss the first draft of the CEGP	Fully
2nd draft of the CEGP reflecting the relevant aspects from the participatory process with stakeholders	Fully
Actions to promote RRI key Integrity	Partially
Actions to publicize the idea of ethical governance of R&I in line with the ETHNA System	Partially (needs continued effort)
Actions to raise internal awareness concerning the Code of Ethics and Good Practices	Partially (needs continued effort)
Actions to promote RRI Training for ESR	Partially (needs continued effort)
Designation of RRI Officer	Not achieved
Established the core duties of RRI Office	Partially
Action Plan for the implementation of the RRI Office	Partially
Actions to promote RRI key Gender	Fully
Actions to issue reports and make recommendations on principles related to conflicts of interest	Partially
Actions to promote RRI key Open Access	Fully (although needing continuous update)

Performance Indicator	Quantification (if applicable)
The level of commitment to ETHNA System determined	Level 2
Tackle the RRI keys: Research Integrity, Gender Perspective, Open Access and Public engagement	Public engagement was achieved only in a limited way
Actions aimed at reflecting, reporting, and making recommendations on principles related to R&I ethics and professional ethics	Multiple actions taken (e.g., interviews + workshops + internal sessions
Actions to monitor the level of compliance by professionals and by the organization with the CEGP values, principles, and behaviors	Only achieved to a limited extent. Further development needed
Actions implemented to generate internal awareness of the ETHNA System	Dissemination through digital channels + Multiple Internal meetings
Actions implemented to generate internal awareness of the contents of the CEGP and its benefits	
Actions CTS has taken to promote the CEGP	
Actions to promote RRI Training	Multiple education actions
Meetings for the creation of RRI task force	Multiple internal meetings
Establishment of the RRI core duties and Action Plan	
Actions taken to promote gender balance	Specific EU project Proposal + Conference paper + Gender and Inclusion Equality Plan + Internal dissemination
Meetings for the creation of the code of conflicts of interest	Multiple internal meetings
Actions taken to promote RRI key Open Access	

4 Implementation Barriers and Drivers

During the implementation process, the following main barriers and drivers were identified:

Barriers. Some of the main barriers found along the implementation are:

- The living lab process proposed by ETHNA was perceived as a bit too bureaucratic/complex, giving implementors the feeling of doing things "just because …". In fact, the method was like a "one way" approach, leaving little room for adaptation to each context. A co-creation approach [17] would better suit the needs, and as such we had to make several changes to the process.
- A major difficulty was identifying external stakeholders and finding ways to involve them. Broadening the conversations on RRI to include wider stakeholder groups typically "conflicts" with more traditional views of scientific research [22]. The situation was made even more difficult due to the COVID-19's confinement period.
- The design of the ETHNA System was, to some extent, based on the assumption that "stakeholders are eager to adopt these ideas and get involved." This is not really the case, as researchers are busy with their own projects and are not so thirsty to spend time on other activities. In this context, one of the main challenges is to find ways to motivate them and design mechanisms that can lead to a "change in culture". The initial ETHNA's guidelines were not very helpful in this change process.

Drivers. On the other hand, a number of factors played a driving role:

- Implementation was greatly facilitated by the fact that the Director of UNINOVA-CTS was involved in the process from the beginning and is the coordinator of the RRI task force.
- The identified actions were considered to contribute to the fulfilment of UNINOVA-CTS' commitments to the Portuguese research funding agency regarding the implementation of a RRI model. Since the outcome of such commitments must be demonstrated during the evaluation of all national centers by that agency, this led to a strong additional motivation.
- The fact that most employers of UNINOVA-CTS researchers (e.g., NOVA University of Lisbon, Polytechnic Institutes, etc.) have already established some level of awareness on some aspects related to RRI, having included in their governance structures ethics committees and mechanisms for promoting good research and innovation practices, gender and inclusion plans, etc., facilitated the initial dialogue.
- Most UNINOVA-CTS researchers are members of international and national scientific and technical societies (e.g., IEEE, IFIP, IFAC, Socolnet, National Engineers Association) and, as such, they are asked to comply with the code of ethics of such associations. Nevertheless, in practice, many of them do not pay much attention to those codes.
- The proposed ETHNA System implementation methodology, which is carefully designed, well detailed, documented and explained, despite the lack of co-creation mechanisms as mentioned above, was useful in giving a broad overview of the process.

Based on the acquired experience, some changes to the general ETHNA System guidelines could be perceived as good practices:

- Starting with an additional action to prepare first drafts of key documents by RRI task force, rather than multiplying the number of workshops that require the involvement of the various stakeholders. Only after these draft documents are produced should we proceed with consultation and refinement by involving all relevant stakeholders in the organization as a way to engage them in a participatory process.
- Making an effort to adapt the "jargon" of the ETHNA System to the internal reality and constraints of the center. This also requires extra effort from the RRI task force but contributes to better acceptance by the stakeholders.
- Establishing a strong and clear link between the RRI implementation process and the internal preparation for the center's evaluation by the national funding agency (which calms potential internal "political" barriers).

5 Lessons Learned

From the experience acquired with this implementation exercise, a number of learned lessons can be mentioned:

On the Sustainability of Institutional Changes. An important aspect in any RRI implementation project is how to make the initiated actions sustainable in the long term. Although it is too early to make an assessment of such sustainability, some points can be mentioned:

– Considering the reality of the center, it was understood that the most important and long-lasting measure is the change of culture. This is rather difficult to achieve with older/senior researchers who have their own habits and autonomy. Senior researchers are also too busy with their own research and management activities. Therefore, the emphasis is put on the training young researchers. The center hosts about 100 PhD students and they all receive specific training on research ethics and RRI, which is expected to be an effective mechanism for changing the culture. But this needs to be a continued effort, as we receive new students every year.
– On a more "political level", the tools provided by the ETHNA System greatly helped UNINOVA-CTS to fulfil its contractual obligations to the national research funding agency, which would otherwise require too long a process with long discussions and too many "political" hurdles.

On the Wider Potential of Institutional Changes at the Organization. Beyond the internal impacts of an RRI implementation, namely in terms of mechanisms and change of behaviour, it is also relevant to consider the wider impact in terms of relationships with external stakeholders:

– Prior to the implementation of ETHNA, the center had no experience or even awareness of discussing RRI with external stakeholders. As a result of the implementation

exercise, several contacts and working sessions took place with external stakeholders. The overall reaction was quite positive and good feedback was obtained to improve our key RRI documents/plans.
– These interactions opened some initial directions for further collaboration with those external entities (new project proposal, joint conference paper, etc.). In fact, this is also in line with the growing awareness of the social responsibility of research institutions regarding technology transfer [2]. An increased perception of the role of co-creation [17] also results from these interactions.

On What Worked Well. Several aspects of the implementation process proved to be particularly adequate for a networked ecosystem like our center. Among these we can highlight:

– Adopting a "complementarity" perspective, i.e., focusing only on aspects deemed important to the center (understood as a federated ecosystem of researchers), and assuming that other general ethical principles are already covered by the ethical codes of the different employing institutions with which our researchers have a contractual link, is an effective approach.
– Avoiding the creation of complex bureaucratic organizational structures. Since there are no funds for such structures, having only a task force was manageable at this stage. However, it is still too early to assess whether it works in the long term.
– Rather than attempting to elaborate the key documents through a large number of discussion workshops (as suggested in the initial living lab methodology), we created a small team that did substantial preparatory work and drafted preliminary versions of these documents. Then the internal workshops were minimized and focused on discussing these drafts and collecting feedback, which proved to be an effective approach**.**

On What was Difficult. In addition to the effort to smooth internal "political barriers" and to interact with external stakeholders, other aspects of the process proved particularly difficult. Among these we can highlight:

– As mentioned before, adopting the "living lab guidelines" of ETHNA appeared a bit too "bureaucratic" for our reality. The implementation process and corresponding guidelines should be a co-creation process, rather than having to follow a top-down prescription. The model recommended in the project was too "unidirectional", from "designers to implementers", which is very wrong and against the very essence of a living lab where co-creation should be nurtured.
– Another thing that perhaps happens with projects focused on RRI implementation, is that at the beginning, part of the consortium had the assumption that "Stakeholders are eager to adopt these ideas and get engaged". This is not really the case, as researchers in our institutions are busy with their own projects and while they may see RRI as important, they are not available to spend much time on a complex implementation process or to accept more bureaucratic procedures.
– As also mentioned above, the most important thing is to "change the culture" and the current living lab methodology needs further refinements to support this.

Recommendations. Although this work is based on a single implementation experience, we believe that some recommendations might be useful for similar research ecosystems:

- When starting any RRI implementation process, it is necessary to make a good demonstration of the importance of RRI and the potential benefits (added value) to all participants. This can be challenging since we are competing for the time and attention of researchers (and other stakeholders) that are focused on their own activities.
- Organize a small RRI team that takes the ETHNA System tools and examples and makes an adaptation to the reality of the organization. This adaptation will make the process much more effective than starting from scratch.
- It is crucial to involve the organization's leadership in the process.
- If possible, link closely to strategic commitments between the organization and its relationship with national funding agencies. As such, participants feel the pressure from outside rather than from an internal taskforce.
- Whenever possible, establish a link with training programs for young researchers in order to create a new culture. This seems to be in line with some emerging trends for teaching RRI in higher education institutions [23, 24].
- Considering the current ETHNA System living lab guidelines, it is necessary to simplify and adapt them to the reality of each organization.

Naturally the implementation of the ETHNA System [19] is not a "one shot" process, especially in the context of a dynamic (and volatile) research and innovation ecosystem. The process needs continuous attention and monitoring [21] as societal demands evolve and, in the case of an ecosystem like UNINOVA-CTS, there is a continuous flow of people in/out.

6 Conclusions

Effectively introducing RRI dimensions and key mechanisms into engineering and technology-oriented research ecosystems is not easy, because there is less awareness in these communities of the importance of the principles at stake compared to other communities, such as health and biomedical research. Researchers in engineering and technology development are often too busy with their own projects and attracting new funded projects, being difficult to engage them in discussing other issues. Therefore, raising awareness and implementing mechanisms for culture change are major challenges. However, the adoption of a systematic method as proposed by ETHNA and using a collection of examples and templates greatly facilitates the process.

In the case of a research and innovation ecosystem, it is important to adopt a "complementarity-based" approach, as participants are subject to various "RRI spaces", each one with specific requirements. This is particularly relevant as more and more research is carried out by dynamic networked communities.

As the number of researchers focused on RRI increases, namely in the area of social sciences, and new focused projects are launched, the area consolidates, particularly in

terms of concepts, mechanisms, and procedures, but there is a risk that this "new community" becomes too separated from the "other researchers" for whom these developments are made. Thus, there is a need for a continuous effort of co-creation and adaptation to the reality of each organization.

This work reflects an initial implementation process, lacking the long-term feedback. Continuous monitoring and adjustments are necessary as societal perspectives and demands on RRI evolve.

Acknowledgments. We acknowledge the support from the European Commission through the ETHNA System project (Grant Nº 872360). Partial financial support was also provided by Portuguese FCT program UIDB/00066/2020 (CTS – Center of Technology and Systems).

References

1. Burget, M., Bardone, E., Pedaste, M.: Definitions and conceptual dimensions of responsible research and innovation: a literature review. Sci. Eng. Ethics **23**(1), 1–19 (2016). https://doi.org/10.1007/s11948-016-9782-1
2. Gwizdała, J., Śledzik, K.: Responsible research and innovation in the context of university technology transfer. Folia Oeconomica Acta Universitas Lodzensis **2**(328) (2017). https://doi.org/10.18778/0208-6018.328.04
3. Liu, J., Zhang, G., Lv, X., Li, J.: Discovering the landscape and evolution of responsible research and innovation (RRI): science mapping based on bibliometric analysis. Sustainability **14**, 8944 (2022). https://doi.org/10.3390/su14148944
4. Bouter, L.: What research institutions can do to foster research integrity. Sci. Eng. Ethics **26**(4), 2363–2369 (2020). https://doi.org/10.1007/s11948-020-00178-5
5. United Nations: Transforming our world: the 2030 agenda for sustainable development. Department of Economic and Social Affairs. https://sdgs.un.org/2030agenda (2015). Accessed 6 Apr 2021
6. Huang, C., Zhang, Z., Mao B. and Yao, X.: An overview of artificial intelligence ethics. IEEE Trans. Artif. Intell. (2022). https://doi.org/10.1109/TAI.2022.3194503
7. Sabel, B.A., Seifert, R.: How criminal science publishing gangs damage the genesis of knowledge and technology—a call to action to restore trust. Naunyn Schmiedebergs Arch. Pharmacol. **394**(11), 2147–2151 (2021). https://doi.org/10.1007/s00210-021-02158-3
8. Jiya, T.: Responsible data governance in projects: applying a responsible research and innovation (RRI) framework. J. Technol. Manage. Innov. **16**(1) (2021). https://doi.org/10.4067/S0718-27242021000100031
9. Camarinha-Matos, L.M., Afsarmanesh, H.: Collaborative networks: a new scientific discipline. J. Intell. Manufact. **16**(4–5), 439–452 (2005). https://doi.org/10.1007/s10845-005-1656-3
10. Graça, P., Camarinha-Matos, L.M.: Performance indicators for collaborative business ecosystems – literature review and trends. Technol. Forecast. Soc. Change **116**, 237–255 (2017). https://doi.org/10.1016/j.techfore.2016.10.012
11. Camarinha-Matos, L.M., Fornasiero, R., Ramezani, J., Ferrada, F.: Collaborative networks: a pillar of digital transformation. Appl. Sci. **9**(24), 5431 (2019). https://doi.org/10.3390/app9245431

12. Mehari, Y., Pekkola, E., Hjelt, J., Cai, Y., Stenvall, J., Ortega-Colomer, F.J.: Defining 'responsible' in responsible research and innovation: the case of quadruple helix innovation in the energy sector in the tampere region. In: Păunescu, C., Lepik, KL., Spencer, N. (eds.) Social Innovation in Higher Education. Innovation, Technology, and Knowledge Management, pp. 199–225. Springer, Cham. (2022). https://doi.org/10.1007/978-3-030-84044-0_10
13. Nieminen, M., Ikonen, V.: A future- oriented evaluation and development model for Responsible Research and Innovation. In: Assessment of Responsible Innovation - Methods and Practices, Taylor & Francis, New York (2021). https://library.oapen.org/bitstream/handle/20.500.12657/42836/9781000292749.pdf?sequence=1#page=275
14. Forsberg, EM., Shelley-Egan, C., Ladikas, M., Owen, R.: Implementing Responsible Research and Innovation in Research Funding and Research Conducting Organisations—What Have We Learned so Far?. In: Governance and Sustainability of Responsible Research and Innovation Processes. SpringerBriefs in Research and Innovation Governance. Springer, Cham (2018)
15. EC: The European Charter for Researchers. European Commission, Directorate-General for Research (2005). https://euraxess.ec.europa.eu/sites/default/files/am509774cee_en_e4.pdf. Accessed 28 Jan 2023
16. IEEE: Definition of authorship. https://journals.ieeeauthorcenter.ieee.org/become-an-ieee-journal-author/publishing-ethics/ethical-requirements/. Accessed 5 Jan 2023
17. Robinson, D.K.R., Simone, A., Mazzonetto, M.: RRI legacies: co-creation for responsible, equitable and fair innovation in Horizon Europe. J. Respons. Innov. **8**(2), 209–216 (2021). https://doi.org/10.1080/23299460.2020.1842633
18. Griessler, E., Braun, R., Wicher, M., Yorulmaz, M.: The drama of responsible research and innovation: the ups and downs of a policy concept. In: Blok, V. (ed.) Putting Responsible Research and Innovation into Practice. Library of Ethics and Applied Philosophy, vol. 40, pp. 11–34. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-14710-4_2
19. González-Esteban, E.: The ETHNA System and Support Tools. In: Ethics and Responsible Research and Innovation in Practice, LNCS, volume 13875, Springer, Cham (2023)
20. Camarinha-Matos, L. M., Goes, J., Gomes, L., Pereira, P.: Soft and transferable skills acquisition through organizing a doctoral conference. Educ. Sci. **10**(9), 235 (2020) https://doi.org/10.3390/educsci10090235
21. Monsonís-Payá, I., García-Melón, M., Lozano, J.-F.: Indicators for responsible research and innovation: a methodological proposal for context-based weighting. Sustainability **9**(12), 2168 (2017). https://doi.org/10.3390/su9122168
22. Stahl, B.C., Obach, M., Yaghmaei, E., Ikonen, V., Chatfield, K., Brem, A.: The responsible research and innovation (RRI) maturity model: linking theory and practice. Sustainability **9**(6), 1036 (2017). https://doi.org/10.3390/su9061036
23. Mejlgaard, N., et al.: Teaching responsible research and innovation: a phronetic perspective. Sci. Eng. Ethics **25**(2), 597–615 (2018). https://doi.org/10.1007/s11948-018-0029-1
24. Lukovics, M., Udvari, B., Nádas, N., Fisher, E.: Raising awareness of researchers-in-the-making toward responsible research and innovation. J. Knowl. Econ. **10**(4), 1558–1577 (2019). https://doi.org/10.1007/s13132-019-00624-1

Developing RRI and Research Ethics in Universities

Laura Bernal-Sánchez(✉) and Ramón A. Feenstra

Universitat Jaume I, Avenida SosBaynat, s/n, 12071 Castelló, Spain
{lbernal,feenstra}@uji.es

Abstract. This work examines the process of developing responsible research and innovation (RRI) at Universitat Jaume I, a public university in Castelló Spain. In this context, the chapter presents some basic characteristics of RRI implementation by exploring thestepsin the process, and some barriers and drivers. In particular, the authors examine the development of a code of good research practice and the university's ethics committees. Both tools are developed in line with the European ETHNA System project, which provides practical guides for RRI institutionalisation processes.

Keywords: Research ethics · ethical tools · code of good practices · ethics committees · university · Universitat Jaume I

1 Research Ethics and Integrity in Universities: Why is It a Crucial Topic?

Research ethics is an issue of vital interest for universities, as institutions whose mission combines teaching and research [1]. Research offers endless and unquestionable benefits when it is conducted in an honest and reputable manner. It advances scientific knowledge, thereby contributing to social improvement and/or wellbeing by helping to solve all manner of problems, and it ensures educational curricula keep abreast with the most recent knowledge, among many other aspects.

Nevertheless, research is not always performed according to the standards and aims that should be expected [2–5].[1] For instance, the basic principles protecting research participants' personal data are sometimes violated, or informed consents are not managed correctly. On other occasions, research misconduct takes the form of duplicate publications, failure to cite research work that should be mentioned or, conversely, including citations to boost the impact of colleagues' or friends' research. On the other hand, relationships with colleagues, especially those in the earlier stages of their career, may be subject to manipulation or abuse of power. Such bad practices are a growing concern,

[1] For a reference document on good practices and research misconduct, see The European Code of Conduct for Research Integrity. https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf [6].

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 109–121, 2023.
https://doi.org/10.1007/978-3-031-33177-0_7

particularly because they proliferate in fiercely competitive and highly demanding contexts like universities [7–9]. It is therefore unsurprising that the concept of "publish or perish" is often criticised as being potentially harmful to scientific integrity [10].

Research ethics and the governance of ethics (see Chapters 1 and 2 in this volume), particularly the implementation of instruments such as codes of good practices or ethics committees, are especially important in this context. Indeed, they are of paramount importance to promote a scientific culture based on integrity or an honest scientific culture that guides both individual and collective actions [11–13].

The objective of this chapter is to present some of the steps followed in setting up an ethical governance system with RRI tools at Universitat Jaume I (UJI).[2] This process was carried out between 2020 and2022 in the frame of the European ETHNA System project [1, 14].

2 Developing RRI at the UJI

Codes of good practices and ethics committees are two of the main building blocks in processes to implement responsible research and innovation (RRI) [14–18]. They can help to promote transparency in the research process with the aim of aligning research with society's values and priorities, or addressing socio-ethical issues.

Institutionalisation and implementation processes in universities are notoriously problematic due to their complex organisational structure, in which very different and interrelated tasks overlap and complement one another: management, teaching, research and knowledge transfer to society, public engagement, and accountability [19–22]. Universities are divided into decentralised structures and departments that are not free from conflicts and internal tensions [23]. The challenges that arise from implementing RRI in universities have led to calls in the literature for reflection on the meaning of good governance [24–28]. Likewise, attempts are also being made to properly define the steps to follow when implementing RRI, especially given the wide range of aspects involved: roles and profiles, tasks, levels of engagement, regulations and procedures, and the specific structures that support them [29, 30].

During the implementation process at the Universitat Jaume I, every effort was made to bear these aspects in mind and to be aware of the limitations and difficulties involved when promoting research ethics. The process focused especially, albeit not exclusively, on two key instruments: the code of good practices in research and research ethics committees.[3] The implementation took place during a research period of approximately two years and involved three phases: 1) *reviewing* the existing tools in the institution and researchers' perception of ethical research matters and RRI; 2) *participatory* drafting of the code of good practices and the regulations for the new ethics committees; 3) *public presentation*, *debate* and *approval* of documents.

[2] Universitat Jaume I (UJI) is a public university, founded in 1991, in the city of Castelló de la Plana in the Valencian Community, Spain. In 2022, it had a teaching and research staff of 1,363,649 administrative and service staff, and 14,080 students.

[3] Three research ethics committees were set up. The Research Ethics and Integrity Committee (CEI), the Human Research Ethics Committee (CEISH) and the Animal Experimentation Ethics Committee (CEEA).

2.1 The Review

The first step in our implementation process was to observe what other institutions with similar characteristics were doing. Ethics committees and codes of good practice in research at other Spanish universities and research organisations were closely analysed. At the same time, the ETHNA guidelines proved useful for defining potential content for the code of good practice [31]. In addition, we carried out an in-depth analysis of the university's existing structures, regulations, procedures and tools related to research ethics. This process showed that some basic tools were already in place, such as a code of ethics, a deontology committee and an ethics hotline. At the same time, various departments had considerable experience in managing some RRI keys, such as open access (Library Service), gender equality (Equality Unit) and research integrity (the Deontology Commission and the Animal Welfare Ethics Committee). This exploration identified the way the existing ethical governance mechanisms were structured and interrelated.

An internal survey was also conducted between 5 May and 13 June, 2021,[4]which asked the university community about its perceptions and concerns in relation to five thematic areas. Specifically, there were eight questions on open access, four on gender equality, seven on bad research practices, four on knowledge of research ethics, and six on the institution's ethical governance. The survey yielded 539 responses from a total population of 1,030 teachers, representing a response rate of 52.33%.The survey population included researchers from four professional categories (R1, R2, R3 and R4),[5] corresponding to the following levels: R1 (first stage researcher, up to the point of PhD); R2 (PhD holders or equivalent who are not yet fully independent); R3 (established researchers who have developed a certain level of independence); R4 (leading researchers at the top of their research area or field).

The survey covered all areas of knowledge: Arts and Humanities, Health Sciences, Sciences, Social Sciences and Law, and Engineering and Architecture. The initial review phase allowed us to draw a series of conclusions for the implementation process:

- The university had a series of basic RRI tools. Therefore, the implementation project did not need to start from scratch.
- Different units and researchers dealt with essential matters of research ethics, and also of RRI in general, but their tasks were not always coordinated or delimited.
- The scientific community showed an interest in these matters as indicated by the high survey response rate and the fact that researchers provided exhaustive qualitative-type replies to the open question.
- The respondents' knowledge of the existing ethical governance structures was moderate or poor: 41.93% was aware of the university's code of ethics, 30.24% was aware of the Deontology Committee and the Animal Welfare Ethics Committee, and 2.60%knew about the ethics line.

[4] The results of this survey (in Valencian) can be consulted at: https://ujiapps.uji.es/ade/rest/storage/DXN9VMDGG6N03LVYTWCYMOSTGPF23WPO?url=/serveis/opaq/base/gestio-qualitat/mesurament-resultats/informeeticainvestigacion_21.pdf [32].

[5] The population was grouped according to professional category as defined by the European Commission. See: https://euraxess.ec.europa.eu/europe/career-development/training-researchers/research-profiles-descriptors.

- Regarding the question about the prevalence of research misconduct in their knowledge area, 80.15% of the respondents expressed concern about its proliferation at the national level. Some forms of research misconduct were highlighted, such as using personal influence for personal gain (41%) or the abuse of power over researchers in lower positions (26.53%). The least frequent types of misconduct, according to the survey participants, were conflict of interest (6.83%) and plagiarism (2.60%).
- Researchers expressed strong interest in key RRI matters such as open access (73.29%), which they considered important or very important, but acknowledged they knew very little about it: 35.06% of respondents stated that their knowledge on available open access channels was "null" and 34.51% said it was "poor". Similar results emerged for other RRI keys like gender or integrity, albeit to a lesser extent. In general, the data suggested that while there was interest in the RRI keys, respondents knew much less about how to implement or develop them.

2.2 Defining the Implementation as a Participatory Process

The information obtained in the review phase allowed us to define the short- and long-term objectives for the implementation process. Particularly urgent targets were to develop a new code of good research practices, renew the Deontology Committee (and turn it into an ethics committee) and incorporate administration staff members who would be tasked with supporting the research community in matters of research ethics. Similarly, the importance of adopting new tools to address the concerns raised in the survey was noted. One priority in this direction was to provide tools that allowed researchers to integrate key aspects of RRI: gender equality, open access and research integrity, and good (and bad) practices.

Having collected this information and defined the main objectives, a participatory implementation process was begun, following the ETHNA System guide. The aim of this process was to involve a range of stakeholders, especially internal agents who already played a role in matters of research ethics.

In line with the above, the following steps were undertaken:

Step 1. The ETHNA project team members drafted an initial proposal for each of the RRI keys—governance, communication and public engagement, integrity, open access and gender—to be included in the code of good practices and in the regulations of the new ethics committees.
Step 2. These five drafts were discussed with the university's top management in order to ensure that the proposals were aligned with the university's policies.
Step 3. Each draft was debated in five working groups (between January and April 2022) made up of internal stakeholders, each group working on one of the RRI keys. In addition, a working group of external stakeholders was set up. The working groups met in person to discuss the drafts of the code of good practices and the regulations for the new ethics committees.
Step 4. Participants from each RRI key working group were given ten days to comment and make proposals for improvement.

Taking part in this process were 43 UJI community members[6] and seven external members (Table 1). Bilateral meetings were also organised to deal with specific topics, such as data protection, occupational health, and safety.[7]

Table 1. Number of participants in the participatory process

RRI Key or tool	Number of participants in the working groups	Internal/External Stakeholders
Gender	10	Internal
Public Engagement	8	Internal
Open Access	12	Internal
Integrity	10	Internal
Ethical governance	8	Internal
Code of Good Practices	7	External

The process so far can be defined as successful in terms of the level of involvement of participants in each of the working groups. A 'neutral facilitator' figure was appointed to each group, which helped to encourage high levels of participant involvement. This person was familiar with the document under discussion and the whole process, but had not been involved in writing the draft. By proceeding in this way, the neutral facilitator helped to guarantee an open space for debate and ensured that participants would not be influenced by the opinions of the authors of the drafts. The literature on participatory processes strongly advises creating a space for group work in which the participants really feel that their contributions are relevant and are being considered [33, 34].

Another key feature of this participatory process entailed mapping people and services with expertise on each RRI key, which encourages their engagement and favours the efficiency of their interventions.

2.3 Process of Public Presentation, Debate and Approving Documents

Once the documents on the RRI matters had been prepared with the participation of the key stakeholders from the institution, the process opened out to the whole university community. Documents were posted on the university's official website, where all members of the university community could add comments and proposals over a period of several weeks (September 2022).

The documents were then also discussed with representatives of the two main university bodies involved in research: the Commission for Research and Doctoral Studies

[6] Participants who contributed to more than one working group were only counted once in the total figure.

[7] The IT Innovation and Auditing Office (OIATI), the Health, Safety and Environmental Management Office and the Office of the General Secretary of the UJI participated in these meetings.

and the Governing Council. Although only five contributions were received during the public presentation phase, the drafts were once again discussed in the representative bodies. These discussions focused on the definition of certain bad practices, such as duplicate publication; the delimitation of criteria defining conflict of interest; concerns about so-called predatory publishing; and different theoretical and practical approaches to gender mainstreaming in research.

3 Most Outstanding Results

The outcomes of the implementation process were:

1. The Code of Good Practices in Research and Doctoral Studies;
2. New ethical governance structures: the Research Ethics and Integrity Committee, the Human Research Ethics Committee and the Animal Experimentation Ethics Committee;
3. Training and dissemination material on RRI;
4. Appointment of a specialist in research ethics to support the research community (defined as an RRI Officer).[8]

In more detail:

1. The Code of Good Research Practices is structured according to the four main RRI themes. Each section includes commitments and good practices at both institutional (university) and individual (researcher) levels. All these sections also contain a glossary of the basic concepts to clarify understanding of the issues involved (especially among those who recognised their lack of knowledge in the survey).[9]
2. The role of the new Research Ethics and Integrity Committee merits further attention. This committee was created with the purpose of promoting integrity and excellence in research (e.g. developing good practice guidelines). Its structure combines new specialist profiles (public engagement, research ethics, and gender) with existing profiles from the university's top management (head of the research service, head of the doctoral school, vice-rector of research). Some new profiles were also included in the Human Research Ethics Committee and the Animal Experimentation Ethics Committee (specialists in methodology, gender perspective).
3. In addition to the code and the ethics committees, the project has raised awareness among the academic community through new training material. At the same time, the implementation project led to significant improvements in all the information and training material available on the research ethics section of the website (new and more detailed templates for the preparation of the ethics report, models of informed

[8] The conceptualisation proposed by ETHNA appears here (p. 24). https://ethnasystem.eu/wpcontent/uploads/2022/07/D4.2_ETHNA_2022_guide_220210_incl_oolbox_neu.pdf [31].

[9] The dossier with the key concepts: gender equality (p. 14), open access (p. 16) and communication and public engagement (p. 20) can be consulted at: https://ujiapps.uji.es/ade/rest/storage/JC0VO1VPDBAYIPKUZROLFYLCNXLBKD3B?url=/investigacio/base/etica/cbpid/CBPID_cas_ok_web.pdf [35].

consent, FAQs on ethics and research, etc.).[10] For good reason, the literature emphasises training as one of the key points for implementing RRI [37]. The objective of this training is to raise awareness about the relevance of these tools and their usefulness in encouraging the integration of research ethics.[11]

4. Finally, the most relevant change was the incorporation of an RRI Officer. This new professional position helps to ensure an integrated structure within the ethics committees. The RRI Officer has also become the institution's main contact person for advising the research community on the basic principles of RRI.

4 Barriers and Drivers in the Implementation Process

Implementation processes require a series of conditions to be able to make institutional changes, which is normally a complex matter. In this process, drivers and barriers are likely to arise, and can either favour or slow down the process. It is important to be aware of these conditioning factors when promoting research ethics and RRI tools.

Barriers 1 and 2. Some of the most notable barriers in the implementation process, and that tend to be common in other experiences, affect aspects such as lack of knowledge across the researcher community about the main RRI keys, scepticism about the relevance or need to promote research ethics, or the fear of increased research "red tape" [39–43].

The review phase revealed that such problems also affect the UJI. Furthermore, the implementation process showed us that the process demands considerable effort from the participants, especially those tasked with writing and updating drafts after each phase of the participation. The process revealed that, in order to overcome this barrier, the community must be open to participation, and a small, but extremely active, group is needed throughout the implementation process. However,certain drivers allowed us to advance and overcome difficulties during the process. In the case of the UJI, the main drivers fell into two types: those typically associated with the institution and external drivers.

Driver 1. The institutional drivers firstly include the wholehearted commitment of both the university and its top management. Engagement in the governance process takes place at various levels: at the leadership level, at the management level and at the level of the researchers themselves. This bottom-up and top-down involvement favours the engagement of key actors, those with expertise in RRI core issues [44–46]. This first driver is complemented by a series of external drivers, which can perhaps be defined as a general ecosystem that favours the development of research ethics and RRI in general. Such aspects are essential for promoting implementation.

Driver 2. In the specific case of the UJI, the implementation process coincided with the university's commitment to the quality accreditation process of HR Excellence in

[10] Information about tools, procedures, templates, FAQs, etc. can be consulted on the website https://www.uji.es/investigacio/base/etica/ [36].

[11] Training material can be consulted on the website (tools, procedures, templates, FAQs, etc.) https://www.uji.es/investigacio/base/etica/jornades/ [38].

Research (HRS4R) and the principles of the European Charter [47, 48]. This policy testifies to the strong, clear commitment from the top management team, as well as the capacity to mobilise resources, for instance, to hire an RRI Officer.

Driver 3. A further series of external favourable developments derive from the new requirements placed on funders of research and academic journals. Spanish Law 14/2011 on Science, Technology and Innovation highlights the importance of the professional ethics dimension in carrying out research, the role of ethics committees, and the promotion of codes of good practices and the main RRI keys. This law was reinforced with the draft bill of 17 January 2022, which amends Law 14/2011, and establishes the importance of monitoring aspects of ethics and integrity in scientific practices and research work.

Driver 4. Research funders are also promoting the key role of ethics when funding projects. They expect key requirements to be met, such as applying the 3Rs in animal research (replacing, reducing, refining), opening access to results, protecting the integrity and the data of those participating in research with human beings, and taking into account environmental impact and biosecurity parameters, particularly in research with biological agents and genetically modified organisms. These matters are gradually being included in and expected from applicants in European (Horizon 2020, Erasmus +, ERC, Erasmus Mundus, KA2, etc.) and national calls, both public (state research agency calls, FECYT, UJISABIO, Generalitat Valenciana calls, etc.) and private (Foundations such as BBVA, Mapfre, la Caixa, etc.).

Driver 5. Finally, research journals are increasingly requesting an ethics validation report for research proposals, which is issued by the research ethics committee or a similar organisation, as a requirement to publish the research results obtained.

All these requirements, which derive from European or Spanish laws, funders or publications, act as allies in the process to promote research ethics and RRI. This is because, on the one hand, they amend the context and the "game rules" in which research is to be conducted and, on the other hand, they act as a direct incentive to expose the relevance of such matters to the research community. Throughout the implementation process and the discussions held, it was of paramount importance to present these innovations deriving from the "research ecosystem", and to evidence their relevance and specific practical nature (especially where reluctance or scepticism about RRI was observed).

5 Some General Conclusions

The implementation process at the UJI shows, as is also stressed in the ETHNA project (see also Chapter 4 of this book) that the success (or failure) of promoting research integrity relies on several key factors. The literature highlights that the institutionalisation of RRI depends on the essential role of good leadership in promoting research ethics [1, 49–52]. In this way, institutional support offers the conditions that will enable the process to be subsequently and successfully carried out. At the same time, the commitment from the research community is also crucial since their engagement and participation will

allow instruments such as the code of good practices and the ethics committees to be fully implemented and used. As Stirling points out, implementation aims to strike a potential balance between the top-down and bottom-up processes [53]. With this balance, and with a defined participative process, the research community comes to trust the process, making it possible to follow in line with integrative and plural dynamics.

Furthermore, implementation processes such as the one at the UJI bring to the surface some very basic issues; these are:

- the support and conviction of the institution's top management will be a determining factor for implementation to be truly deep-rooted
- the combination or balance between the top-down and bottom-up processes is another essential (and complex) matter in implementation processes
- the set of norms that define research, such as funding agencies, publications, academic associations, etc., will be another fundamental piece of the jigsaw. The current situation is favourable for raising awareness about the relevance of promoting research ethics.

Three final reflections emerged in the implementation process that may be of relevance for other implementers or institutions that wish to embark on the same process:

- implementation needs to be understood as a long-term process that takes small steps successfully and makes minor achievements, but is always an on-going and open process
- those participating in the process need plenty of patience and the capacity for dialogue if they wish to convince people of the usefulness of ethics management tools, rather than imposing them
- devising training courses and/or seminars that raise awareness about issues of RRI is a necessary step for tools, codes and committees to become fully relevant.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission. The authors would like to thank the reviewers for their comments. We are also grateful to Nicté García Soria and Mary Savage for reviewing and editing the text.

References

1. Rodríguez-Ponce, E.; Pedraja-Rejas, L.: La gobernanza en las universidades públicas :una perspectiva estratégica desde la experiencia. In: En Contreras, F.G., Gil, E.G., Ortiz, O.O.L.O., Merchán, M.A.H. (eds.) 2021. Nuevas experiencias en gobernanza universitaria, pp. 17–36. Ediciones Usta, Colombia (2021). https://doi.org/10.15332/li.lib.2022.00320
2. Martinson, B., Anderson, M., de Vries, R.: Scientists behaving badly. Nature **435**, 737–738 (2005). https://doi.org/10.1038/435737a

3. Fanelli, D.: How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS ONE **4**, 1–11 (2009). https://doi.org/10.1371/journal.pone.0005738
4. Gross, C.: Scientific misconduct. Annu. Rev. Psychol. **67**, 693–711 (2016). https://doi.org/10.1146/annurev-psych-122414-033437
5. Feenstra, R.A., Delgado López-Cózar, E., Pallarés-Domínguez, D.: Research misconduct in the fields of ethics and philosophy: researchers' perceptions in Spain. Sci. Eng. Ethics **27**(1), 1–21 (2021). https://doi.org/10.1007/s11948-021-00278-w
6. All European Academies. http://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf, ISBN 978–3–00–055767–5
7. Pupovac, V., Prijić-Samaržija, S., Petrovečki, M.: Research misconduct in the Croatian scientific community: a survey assessing the forms and characteristics of research misconduct. SciEng Ethics. **23**, 165–181 (2017). https://doi.org/10.1007/s11948-016-9767-0
8. Felaefel, M., et al.: A cross-sectional survey study to assess prevalence and attitudes regarding research misconduct among investigators in the middle east. J. Acad. Ethics **16**(1), 71–87 (2017). https://doi.org/10.1007/s10805-017-9295-9
9. Maggio, L., Dong, T., Driessen, E., Artino Jr., A.: Factors associated with scientific misconduct and questionable research practices in health professions education. Perspect. Med. Educ. **8**(2), 74–82 (2019). https://doi.org/10.1007/s40037-019-0501-x
10. Rawat, S., & Meena, S.: Publish or perish: Where are we heading? J. Res. Med. Sci. **19**, 87–89 (2014). https://pubmed.ncbi.nlm.nih.gov/24778659/
11. Aranzadi, J.: The possibilities of the acting person within an institutional framework: goods, norms, and virtues. J. Bus. Ethics **99**, 87–100 (2011). https://doi.org/10.1007/s10551-011-0751-y
12. Responsibility navigator. http://responsibility-navigator.eu/wp-content/uploads/2016/01/Res-AGorA_Responsibility_Navigator.pdf
13. Burget, M., Bardone, E., Pedaste, M.: Definitions and conceptual dimensions of responsible research and innovation: a literature review. Sci. Eng. Ethics **23**(1), 1–19 (2016). https://doi.org/10.1007/s11948-016-9782-1
14. Von Schomberg, R.: A vision of responsible research and innovation. In: Owen, J.B.R., Heintz, M. (eds.) Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society 2013, pp. 51–74. Wiley (2013)
15. The Ethna System: A Guide to the Ethical Governance of RRI in Innovation and Research in Research Performing Organisations and Research Funding Organisations. https://ethnasystem.eu/wp-content/uploads/2021/07/guide-to-ethna-system.pdf
16. Felt, U.: Responsible research and innovation. In: Sarah, G., Barbara, P.-S., Stephen, H., Janette, L. (eds.) Handbook of genomics, Health and Society, pp. 1–10. Routledge (2018, preprint)
17. Wittrock, C., Forsberg, E.: Handbook for organisations aimed at strengthening Responsible Research and Innovation. European Commission (2019). https://www.researchgate.net/publication/339390760_Handbook_for_Organisations_Aimed_at_Strengthening_Responsible_Research_and_Innovation_Responsible_Research_and_Innovation_in_Practice_Organisation_responsible_for_the_report
18. Kaplin, W.A., Lee, B.A.: The Law of Higher Education. Wiley, Hoboken (2007)
19. Acosta Silva, A.: Gobernanza y desempeño universitario. Revista Venezolana de Gerencia **23**, 432–440 (2018). ISSN 1315–9984
20. Francisco, G., Armenio, P., Juan, M.: Paradigmas emergentes en la Gobernanza universitaria: una aproximación teórica, Utopía y Praxis Latinoamericana. Redalyc **23**(83), 123–136 (2018). https://www.redalyc.org/articulo.oa?id=27957772012

21. Acosta Silva, A.; Gómez González, S.J: Régimen de políticas, gobernanza y desempeño institucional en las universidades públicas estatales en México. In: Contreras, F.G., Gil, E.G., Ortiz, O.O.L.O., Merchán, M.A.H. (eds.) Nuevas experiencias en gobernanza universitaria 2021. pp. 175–208. Ediciones Usta, Colombia (2021). https://doi.org/10.15332/li.lib.2022.00320
22. Whitley, R.: Changing governance and authority relations in the public sciences. Minerva **49**, 359–385 (2011). https://doi.org/10.1007/s11024-011-9182-2
23. Touriñán López, J.M.: La 'tercera misión' De La Universidad, Transferencia De Conocimiento Y Sociedades Del Conocimiento. Una aproximación desde la pedagogía». Contextos Educativos. Revista De Educación **26**, 41–81 (2020), https://doi.org/10.18172/con.4446
24. Maritza, A.: Modelo ético para el éxito de una gestión universitaria orientada en valores de Responsabilidad Social. Redalyc **28**(69), 517–532(2012). https://www.redalyc.org/articulo.oa?id=31025702005
25. Maria Guadalupe, N., Salom, J., Venancio, R., Annherys, P.: Responsabilidad Social Universitaria: enfoque de gestión ética compartida. Redalyc 28(69), 579–594 (2012). https://www.redalyc.org/articulo.oa?id=31025702007
26. Cardiel, H.C., Gómez, R.R.: Gobierno y gobernanza de la universidad: el debate emergente. Recyt. **66**, 11–14 (2014). https://recyt.fecyt.es/index.php/BORDON/article/view/Bordon.2014.66100
27. Madrid, J.H.C.: Asuntos de gobierno universitario. Editorial Pontificia Universidad Javeriana (2016)
28. Bengoetxea Castro, E.: Int. J. Educ. Technol. High. Educ. **9**(2), 267–279 (2012). https://doi.org/10.7238/rusc.v9i2.1415
29. Maassen, P.: Cambiosen la configuración de la gobernanza: una interpretación de la introducción de nuevas estructuras de gestión en la educación superior. En: Barbara, M., Keh (eds.) La nueva gobernanza de los sistemas universitarios, pp. 99–132. Editorial Octaedro (2012)
30. González-Esteban, E., Feenstra, R., Calvo, P., Sanahuja Sanahuja, R., Fernández Beltrán, F., García-Marzá, D.: The ETHNA System – A Guide to the Ethical Governance of RRI in Innovation and Research in Research Performing Organisations and Research Funding Organisations. Deliverable 4.2 Draft concept of the ETHNA System. ETHNA Project [872360] Horizon 2020. Zenodo (2020). https://doi.org/10.5281/zenodo.6532789
31. Resultats de l´enquesta sobre Ètica de la Investigació (2021). https://ujiapps.uji.es/ade/rest/storage/DXN9VMDGG6N03LVYTWCYMOSTGPF23WPO?url=/serveis/opaq/base/gestio-qualitat/mesurament-resultats/informeeticainvestigacion_21.pdf
32. Asaba, E., Suarez-Balcazar, Y.: Participatory research: a promising approach to promote meaningful engagement. Scand. J. OccupTher. **25**, 309–312 (2018). https://doi.org/10.1080/11038128.2018.1541224
33. Packer-Muti, B.: Conducting a focus group. Qual. Rep. TQR. **15**, 1023–1027 (2010). https://doi.org/10.46743/2160-3715/2010.1199
34. Morgan, D.: Focus groups and social interaction. En: Gubrium, J.F., Holstein, J.A., Marvasti, A.B., McKinney, K.D. (eds.) The SAGE Handbook of Interview Research, 2nd edn. pp.161–176. SAGE (2012)
35. Código de Buenas Prácticas en Investigación y Doctorado. https://ujiapps.uji.es/ade/rest/storage/JC0VO1VPDBAYIPKUZROLFYLCNXLBKD3B?url=/investigacio/base/etica/cbpid/CBPID_cas_ok_web.pdf
36. Research Ethics. https://www.uji.es/investigacio/base/etica/
37. Adler, A.H.: The importance of applied ethics for sustainability and social responsibility. Responsib. Sustainab. ZENODO **7**, 3 (2022). https://doi.org/10.5281/zenodo.6576674
38. Research ethics conference, https://www.uji.es/investigacio/base/etica/jornades/

39. Schmal, R., Cabrales, F.: El desafío de la gobernanza universitaria: el caso chileno. Ensaio: Avaliação e Políticas Públicas em Educação. Ensaio **26**, 822–848 (2018). https://doi.org/10.1590/s0104-40362018002601309
40. Dumville, J.C., Watson, J., Raynor, P., Torgerson, D.J.: Research governance: a barrier to ethical research? QJM **97**, 113–114 (2004). https://doi.org/10.1093/qjmed/hch026
41. Sayers, G.M.: Should research ethics committees be told how to think? BMJ J. **33**, 39–42 (2007). https://doi.org/10.1136/jme.2005.014688
42. Quiroz, E.: Por qué auditar a los comités de ética en investigación. Rev. Perú. med. exp. salud publica. **27,** 443–448 (2010), http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-46342010000300019&lng=es&nrm=iso
43. DuBois, J.M., Antes, A.L.: Five dimensions of research ethics: a stakeholder framework for creating a climate of research integrity. AAMC **93**, 550–555 (2018). https://doi.org/10.1097/ACM.0000000000001966
44. Ethics governance in Scottish universities: How can we do better? A qualitative study. The Mason Institute Blog. https://blogs.ed.ac.uk/mason-institute/2022/06/28/ethics-governance-in-scottish-universities-how-can-we-do-better-by-cristina-douglas-and-edward-dove/
45. Carrizo, L.: Producción de conocimiento y políticas públicas: desafíos de la universidad para la gobernanza democrática. Reencuentro **27**, 69–84 (2004). https://www.researchgate.net/publication/26463208_Produccion_de_conocimiento_y_politicas_publicas_Desafios_de_la_Universidad_para_la_gobernanza_democratica
46. Jung, H.S., Silva, M.W.D., Calderón, A.I., Amaya, W.S., Contreras, F.A.G.: Retos a la gobernanza universitaria: acotaciones sobre la cuestión de la autoridad y la profesionalización de la gestión de las universidades. Fronteiras **8**, 435–456 (2019). https://doi.org/10.21664/2238-8869.2019v8i3
47. The Universitat Jaume I application to acknowledge hrs4r: Action Plan. https://ujiapps.uji.es/ade/rest/storage/JRMIK4WDEORORDR6CMVF48NWC8CVXMT0?url=/investigacio/hrs4r/implementacio/plan/planaccion.pdf
48. Human Resources Strategy for Researchers, HRS4R Technical Guidelines for Institutions. https://euraxess.ec.europa.eu/sites/default/files/policy_library/20220720_technical_guidelines_hrs4r.pdf
49. Pérez, A., Fernández, A.R., de Aguilar, S.H.: Gobernanza universitaria y valores: la función de control en la gestión universitaria. Opción **86**, 176–200 (2018). https://www.researchgate.net/publication/329588972_Revista_Opcion_Gobernanza_universitaria_y_valores_la_funcion_de_control_en_la_gestion_universitaria
50. Moorman, R.H., Darnold, T.C., Priesemuth, M.: Perceived leader integrity: supporting the construct validity and utility of a multi-dimensional measure in two samples. Leadersh. Q. **24**, 427–444 (2013). https://doi.org/10.1016/j.leaqua.2013.02.003
51. Lawton, A., Macaulay, M.: Ethics management and ethical management. In: Ethics and Integrity in Public Administration: Concepts and Cases, pp. 119–132. Routledge (2015)
52. Gaete, H., Huerta, P.: Liderazgo estratégico en educación superior. En: Ganga Contreras, F., Leyva Cordero, O., Hernández Paz, A.A., Tamez González, G. (eds.) Investigaciones sobre gobernanza universitaria y formación ciudadana en Educación. pp. 95–119. Editorial Fontanamara (2019)
53. Stirling, A.: "Opening up" and "closing down" power, participation, and pluralism in the social appraisal of technology. Sci. Technol. Human Values **33**, 262–294 (2008). https://doi.org/10.1177/0162243907311265

Ethics Tools in Practice

Relevance and Challenges of Ethics Committees

Kristi Lõuk(✉)

University of Tartu, Ülikooli 18, 50090 Tartu, Estonia
kristi.louk@ut.ee

Abstract. This article explores the role of ethics committees in ensuring ethical research and innovation, which is essential for maintaining trust in science and innovation. The paper argues that the ethics infrastructure must function at the levels of research and innovation institutions, as well as regional and national levels, and should follow agreed-upon rules and requirements. The changing role of ethics committees is discussed, emphasizing their involvement in investigating potential research misconduct and coordinating the activities of all ethics committees within an institution. The article concludes that for a research and innovation institution, it is of crucial importance to consider the entire ethics infrastructure, including the functions of different ethics committees and ways of implementation through transparency, involvement, policies, procedures, and communication to safeguard that both researchers and the public have trust in ethics infrastructure and ethics committees.

Keywords: ethics committees · research ethics · research integrity · responsible research and innovation · governance · oversight · trust in science

1 Introduction

Science is a social arrangement in which different stakeholders (e.g. research-performing organizations, research funding organizations, publishers, and ethics committees) should fulfill their distinctive roles.

Since science is a collective enterprise, it is necessary to build mutual trust. Trust should be built between researchers and innovators, and with society. Researchers and innovators must be able to trust the results of previously conducted research. Society must be able to trust researchers and innovators by giving them the means to carry out research and contribute to it, e.g. participating as research subjects or participating in co-creation. Researchers and innovators can be trusted if they are trustworthy, meaning that they behave honestly, are objective, respect the autonomy and privacy of research subjects, treat animals and the environment with care, and are responsible. To ensure trust in science, it is necessary not only that researchers and innovators fulfill their responsibilities, but also that they have the willingness to do so. Researchers and innovators and research and innovation organizations (both research performing and funding organizations) bear the responsibility for fostering trust in science and ethical research.

Research and innovation organizations should create and sustain environments that encourage ethical research and innovation through education and clear policies e.g. the

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 125–136, 2023.
https://doi.org/10.1007/978-3-031-33177-0_8

Singapore Statement [1]. In The Bonn PRINTEGER Statement [2] the role of the leaders of institutions is specified and institutional responsibilities are described, e.g. increasing transparency in misconduct cases, having effective and safe whistle-blowing channels, establishing a research integrity committee, and appointing an ombudsperson. Moreover, in its report the European Science Foundation has stressed the need to set structures at the national level as well. They have outlined that successful approaches to promoting good research practices include establishing an adequate institutional framework, which includes research ethics committees and research integrity offices, both at the institutional and national levels [3]. Additionally, it should be pointed out that the governance system of ethical research and innovation should fit each country, taking into account its size, research infrastructure, and available resources [3]. Furthermore, if unintended consequences emerge, the system should be revised or policies should be changed accordingly [4]. Also, it should be analyzed whether the members of the research and innovation institutions are motivated to act according to operating principles.

According to Sarah R. Davies, issues related to research integrity can also be regarded more broadly, as a form of soft governance implemented through codes, norms and ELSI (ethical, legal, and social implications) activities [5]. Therefore, ethics committees should be seen and should function as a part of the governance system of ethical research and innovation. Such a system includes stakeholders (e.g. researchers, administration, case investigators), and procedures and requires an ethics-promoting climate [6]; it also consists of oversight, instruction, and policies [7].

One example of an ethical system of governance for research and innovation in organizations has been developed within the Horizon 2020 project ETHNA. Having a clear implementation and governance system that fits the research-performing or funding organizations including responsible research and innovation (RRI) office(r), ethics codes, policies, and tools shows commitment to ethical research and to carrying out research and innovation activities in a responsible way.

2 Ethical Research: From Fragments to the Whole Research Process

It is expected that research and innovation are done ethically and with high quality. Ethical norms should be followed through all stages of research, from planning and applying for research grants to the publication of research results.

While research ethics has been focusing on the protection of research subjects, research integrity has emerged as a response to misconduct cases, mainly about fabrication, falsification, and plagiarism (FFP). Research misconduct is understood more broadly than FFP, especially in Europe. For example, Foeger and Zimmermann state that "[research misconduct] also covers, for example, destruction of primary data, unjustified authorship, and the sabotage of research activities or dishonest attempts to lower the scientific reputation of another researcher" [8]. A more specific list of possible acts of misconduct is presented by Faintuch and Faintuch [9] comprising the following issues: lack of protection of data, violation of confidentiality, clandestine data access; animal abuse; inadequate human consent regarding aims and benefits, risks and harms,

coercion or exploitation of subjects; ghost authorship[1], questionable personal credits, failure of registration and ethical approval, breach of good (clinical) practices, lack of transparency regarding the relationship with funders, undisclosed commercial or personal interests, noncompliance with publisher ethics, nonprofessional language, salami slicing[2], lack of sharing data with co-investigators, study participants, other authorized parties, harassment, bullying, disrespectful behaviour, dishonest mentoring of students, fellows, junior staff, misuse of research funds, false or exaggerated academic titles, qualifications, professional experience, retaliation against whistle-blowers, deceit, scamps, pseudoscience, malicious misconduct allegations. Often, these practices emerge due to the reward criteria in research that value quantity over quality. This issue needs to be dealt with at the level of science as a system. The aforementioned list includes not only acts from the realm of research integrity but also covers the domain of research ethics. It has increasingly been argued (e.g. Ron Iphofen [13]) that the two of them be viewed together, as both are required for ethical research and innovation. Another definition of research ethics and research integrity is provided by Ana Marušić in this book in her chapter "Evidence-based Research Integrity." Braun, Ravn, and Frankus [14] showed that while research ethics committees do their task before researchers begin the actual research, research integrity offices handle possible misconduct cases after the research has been done and a concrete action or behavior has taken place. Whereas the first of them focuses on planning and design, and the other on conduct and implementation, both are needed to safeguard ethical research. At the practical level, the platform "The Embassy of Good Science" is the outcome of the European Commission's initiated and funded research projects EnTIRE and VIRT2UE within the H2020 program, focusing both on issues of research integrity and research ethics.

3 What Do We Mean When We Talk About Ethics Committees?

When encountering the term 'ethics committee', one must pay attention to its meaning, as any committee or commission that deals with ethical issues can be called an ethics committee. Thus, there are committees that deal with ethical issues in different contexts, with different tasks, and at different levels.

Historically, research ethics committees have the longest tradition in ethical research. Their main task has been to ensure the rights and well-being of human research subjects and/or animals participating in research. As such, they constitute an additional safeguard mechanism, both for subjects and researchers. Based on the written documents submitted to the committee, the task of such committees is to weigh the risks and benefits

[1] Ghost authorship means that someone who has contributed substantially to the manuscript is not named as an author or given acknowledgment. It is problematic as it compromises academic integrity and may mask conflicts of interest [10]. Ghost authorship has also been addressed concerning ethics committees. David Shaw has claimed that when ethics committees improve the quality of a research proposal by suggesting major revisions, they should be acknowledged for their contribution. Ethics committees are ghost authors if they are not credited as authors or their contribution is not mentioned in the acknowledgment section [11].

[2] The term, 'salami slicing' (also salami publication or salami publishing) refers to activities where the smallest publishable unit of data is included in the publication [12].

of the study. Only studies with a proportional relationship between scientific validity, social value, fair participant protection, favorable risk-benefit ratio, independent review, informed consent, respect for participants, and collaborative partnership [15] should be conducted. The role of ethics committees originated in medical research. For example, in the Helsinki Declaration of the World Medical Association [16], it has been a requirement since 1975 to obtain the ethics committee's approval before starting a study. Later, this requirement of prior approval was extended from biomedicine to other fields of research. Social scientists have seen this development as problematic, as the requirement coming from biomedicine may not be suitable for social sciences and humanities [17, 18]. The need to obtain the approval of the ethics committee before starting the research may derive from national legislation, but it may also be required, for example, by research funding agencies or journals where the results of the research are to be published.

At the practical level, the Council of Europe's Steering Committee on Bioethics has worked out a "Guide for Research Ethics Committee Members" [19]. The guide addresses issues related to the appointment of research ethics committee members, emphasizing that the process should be transparent and fair. The requirements for the composition of the committees can be determined either at the level of national law or through guidance documents (e.g. the ETHNA framework) giving recommendations on the number of members of the committee, their qualifications, involvement of lay persons (those whose expertise is not in a specific type of research and who are there to represent the perspective of participants). Having clear policies and making them known is very important for addressing possible conflicts of interest.

However, several critical notes have been addressed to ethics committees. It has been claimed that the requirement of an ethics review is often only formal [20]. From the initial idea to help researchers to balance risks and harms and consider possible benefits of the research and innovation project or proposal [20], the process has boiled down to checking relevant boxes from the checklist or ethics issues table. As such, the "tick the right box" approach does not serve the original purpose of analysing the ethical aspects of one's research and innovation project.

Based on the situation in the United States, it has been claimed that institutional review boards are understaffed, overburdened; that they do not devote enough time, lack sufficient experience, and do not have institutional support [21, 22]. Additionally, there have been discussions on how to define and measure the quality of an institutional review board [23, 24]; how to deal with inconsistencies in research ethics committee review [25]; how to address the disciplinary distinctions in the ethics review process, [26] and the reasons why retrospective review should be added to a prospective one [27].

Although this is quite a tough slate of criticism to be dealt with, these issues can be addressed at the institutional or national level by providing resources, education, and training opportunities. The ETHNA project has given its input by creating a set of guidance tools and a toolbox to help institutions [28]. Additionally, the recommendation has been made that institutional review boards be accredited [29]. Increasing institutional support can mean various things, from hiring more support staff for the ethics committee to providing more training opportunities for members of the institutional review board [21], or requiring ethics committee members to obtain obligatory ethics training.

At the same time, ethics committees can not only review research but also deal with the investigation of possible cases of misconduct. In addition to the 'ethics committee', the body may also bear the name of the 'ethics commission', 'integrity committee', 'research integrity board', or 'misconduct commission', etc. Also, investigations are mentioned in the Bonn PRINTEGER Statement [2], where it is stated that the integrity committee can function either at the institutional or national level. However, the document does not specify whether appeals can be carried out at the national or institutional levels. Whereas the PRINTEGER Statement focuses on the level of research-performing organisations, the ENRIO handbook "Recommendations for the Investigations of Research Misconduct" [30] provides guidance both at the local and national level, e.g. regarding which decisions should be handled at what level and discusses the advantages of local or national bodies. The mutual learning exercise (MLE) in the research integrity final report gives a recommendation "to create a national research integrity body that could help coordinate, monitor, educate, communicate and promote research integrity in a country" [3]. In addition to establishing a national body, having national-level research integrity officers is important [3]. With regard to oversight, the MLE report suggests having an appeal system for research integrity investigations. This would be especially important in countries where there are no research ethics and integrity bodies at a national level [3].

Since the field of responsible research and innovation covers more than research integrity, the ethics infrastructure at all levels should be adapted accordingly. This should go hand in hand with processes determining that necessary regulations and guidance documents on all types of ethics committees are in place. Firstly, the lack of regulations on ethics committees might be a cause for various misconduct cases. Secondly, the ethics infrastructure should cover all aspects of ethics committees, from ethics review to integrity investigations. Thirdly, both should be in place, a regulatory framework as well as good practice guidance documents. We now turn to ethics (governance) bodies at the institutional level of a research and innovation institution.

4 What Should the Ethics Committee(s) Do?

One of the biggest challenges for a research and innovation institution is to work out an ethical governance system that suits the institution and serves it well. For ethics committees within the ETHNA System [28] the following activities are recommended for consideration.

First, the question should be considered what is the purpose of the ethics committee on research and innovation and how does the committee relate to the RRI Office(r) whose task it is to disseminate the concepts of the ETHNA system, establish performance indicators and monitor the progress of the system in the organization? Different options are available, as the committee can function as a governance committee by keeping an eye on the practical implementation of responsible research and innovation, but it can also create the space for discussions over procedures, commitments, and values.

Second, the question arises of the scope of the ethics committee. Who are the internal and external stakeholders to be involved in deliberations on ethical governance? Does the committee cover one or all of the key aspects of responsible research and innovation:

research integrity, gender perspective, open access, and public engagement, or only one of them?

Third, the main principles of action of the ethics committee should be agreed upon. Having such principles, making them known, and acting on them are of crucial importance for building and maintaining trust, both within the institution and with external stakeholders. Within the ETHNA system, the focus is on the following principles: confidentiality, impartiality, fairness, anonymity, the accuracy of information, and fair, respectful, and mediated agreements [28]. One question that arises is whether it is possible to act based on the principle of anonymity when personal contact and additional queries for information might be needed for mediating and agreement.

Fourth, choosing a suitable model of the ethics committee for the research performing or research funding organization is crucial. In a way, this is the cornerstone aspect of the whole system. The ethics committee, its size, needs and resources should fit the institution and should help implementing the ethical governance of RRI in the organization. Does the institution need a standing or an ad hoc committee? Both have their advantages and disadvantages. Having a permanent committee enables continuous operation, so for large research and innovation institutions this would be recommended. However, if the number of situations to be dealt with is low and resources are limited, it would be advisable to have an ad hoc committee instead.

Fifth, what are the functions of the ethics committee members, and what kind of profiles are needed? The way members of ethics committees are selected or nominated and the period of appointment should be known within the institution and by external stakeholders. Other aspects to consider are: how does renewal take place? For how many terms can a person be a member of the ethics committee? How does replacement or dismissal take place? What are the tasks of the secretary? Is the secretary a member of the committee or a member of the committee's support staff? What IT solutions are there or needed to support the work of the committee?

The solution provided by the ETHNA System is to nominate a committee for 4 years, but there are also other possible solutions, e.g. for a period of three years or five years. The next step is to map what profiles of people are needed in the ethics committee. What are the research and innovation and teaching activities the institution is carrying out? For example, doing research with human subjects, animal research, or research on biobanks or genetically modified organisms requires special expertise and potentially also a separate committee for reviewing these research projects. As such, it can be that in an institution there are separate ethics review boards for health research and biobank research, social sciences and humanities, animal research, and genetically modified organisms. One joint characteristic is that all of these committees should have a representative from the Ethics Committee of the institution and an ethics expert, and a law expert. Depending on how you are building up or changing the system, it can also be that the ethics and/or governance committee consists of representatives of the other ethics committees of the institution (mostly chairpersons). For example, the University of Liverpool in the United Kingdom has a research integrity and governance committee [31] that among other tasks has to oversee the work of other committees.

The sixth activity should be about describing the objectives assigned to the ethics committee. Those can include raising awareness, providing advice, helping with conflict resolution, and updating good practice guidelines, among others.

Seventh, what are the principles of the action of the ethics committee? This means that the frequency and interval between committee meetings should be thought through. For standing committees, it can be once a month, and for ad hoc committees whenever necessary. Compared to review and investigation committees, governance committees with the task of oversight and monitoring, and updating policies, should meet a couple of times a year. The schedule should be agreed upon and the information should be made publicly available (on the website of the institution or the webpage of the committee). How is the quorum for decisions decided? In most cases, this means that half of the members should be present. At the practical level, there can be differences, depending on whether or not the secretary is a member of the committee. Do external experts participate in decision-making? Is voting allowed or should deliberations take place until a consensus is reached? How often will the committee issue various reports? How will the body in the institution approve the establishment of the ethics committee on research and innovation?

Eighth, what are the monitoring indicators for the ethics committee? There should be progress indicators and performance indicators and these should be included and made known in the action plan of the ETHNA system.

5 Implementation of the Ethics Infrastructure

For a research and innovation institution, it is important to take into account whether the creation of committees and ethics infrastructure starts from scratch or if the already existing elements are changed and/or developed. If the existing system is changed and developed further, it is necessary to consider, negotiate and communicate tasks and responsibilities. How do ethics committees relate to ombudspersons and/or good science counselors? Would something be taken away from one committee and given to another? Will any new additional committees be created? Will some of the committees have their duties stripped? It is also necessary to consider how all parties affected by these changes are interconnected and how to involve them in the process, providing them with information about possible changes and asking for feedback.

Implementation of the ethics infrastructure also depends on the size of the research and innovation organization or research funding institution as well as on the size and institutional structures of the country where the institution operates. One way to create or change the system is to think about all the different stages of research and innovation and what is needed there. For example, a committee or committees are needed to review the research proposals in the planning stage, then to handle possible misconduct cases. Alternatively, there may also be a need for a committee governing the whole system and providing counseling at all stages of research and innovation. Additionally, does the governance committee also act as an appeal committee, or are there other committees at the regional and/or national level to manage the appeals?

It also needs to be considered whether it is necessary to do everything by yourself in your institution or whether there is a possibility to cooperate, especially in a small

country and in a small research and innovation institution. Additionally, there might be regulatory aspects to be taken into account, such as the ones specifying that some types of research and innovation activities should be reviewed by a specific body not at the same institution, and should rather be done by a regional or national committee(for example this might be the case for clinical trials in some countries). For a small institution in a small country, it can very well be that all the tasks related to governance, review, and investigations take place in one single ethics committee.

To sum up, it is of crucial importance to think through the following questions. Why does the research and innovation institution need an ethics committee? What should the committee do? Do you change the current system or start anew? How do you involve all stakeholders in the process? Do you have the necessary resources (people, infrastructure, etc.) to implement the ethics infrastructure?

When creating a system of ethical governance, it is especially important to avoid making all researchers and innovators feel like their primary task is to constantly prove that all is done by regulations and guidelines of good practice.

Rather than only expecting compliance with codes and regulations, policies should be supportive and aim toward good outcomes. For example, Zwart and Meulen [32] are of the opinion that only using bottom-up processes can foster ethics and integrity in research, and that integrity work must take place in everyday research settings. This should be done by research institutions by facilitating open dialogue and fostering a culture of deliberation, e.g. by creating a safe space for discussing issues from everyday practice. Nevertheless, Zwart and Meulen are of the opinion that the focus should be at the institutional level, not on the level of an individual researcher; not on exposure and punishment, but rather on having a supportive ecosystem. Therefore, a supportive culture is needed, and creating it is the responsibility of the leaders. Furthermore, there should be a variety of training available, for various stakeholders. In this process, not only ethics committees for review and investigation of possible misconduct are needed, but in addition, there should also be persons in the research and innovation institution to whom people can turn to get advice (before handing in an application to get research ethics approval or an application that a misconduct investigation should be started). This can be a good science counselor, ombudsperson, ethics officer, research integrity office or officer. Such persons should also act as mediators giving advice, as well as listening to feedback and the reactions of the users of the ethics infrastructure to further enhance the system.

To implement ethical infrastructure and to help leaders of institutions, another project funded by the European Commission, SOPs4RI, has also worked out templates and guidelines for writing research integrity promotion plans, both for research-performing organizations and research funding organizations [33–35]. The name “promotion plan” might be misleading, since the actual focus is broader, covering preparation (diagnosis, assessing readiness, finding the right people, creating the plan), execution, and monitoring. This is similar to the ETNHA action plan as the plan has to be concrete and contain concrete actions, listing specific responsibilities and deadlines.

At the practical level, it should be thought through how the dissemination of the role and functions of the ethics committees in a research and innovation institution should take place. Furthermore, will there be educational activities provided about the ethical

infrastructure in general for members of different ethics committees (new and current) and for researchers and innovators? Whose task will it be to provide them? The plan should have a concrete bearer of responsibility, be it the ethics officer, the research integrity officer, the ethics committee chair, or someone else.

If the earlier questions remain unspecific, let us consider two concrete examples. First, since the framework program Horizon Europe, the European Commission changed the table on ethical issues by adding the category artificial intelligence [36]. Who in the research and innovation institution should be aware of this requirement and who has to organize educational activities and the communication plan for those to whom it is relevant? Furthermore, does this require that the duties of the ethics review committee members have to be updated and the members educated, so researchers and innovators can submit research protocols encompassing AI to the ethics review committee of the research and innovation institution? Does the institution need to update its policies? For more specific debate about issues related to AI, see the chapter "Ethics and Development of Advanced Technology Systems in Public Administration" by António da Costa Alexandre and Luís Moniz Pereira in this book.

The other case is about chatbot ChatGPT, the AI tool that can create sentences and, in some cases, has been listed as an author of a research paper [37]. Who has to provide information about the possible threats of this development? Should the guidelines and policies about investigating possible misconduct be changed? What can the governance committee do regarding this challenge? Do the institution and ethics committee members have the knowledge to assess whether a text (e.g. research project submitted for ethics review) has been written by a researcher or a chatbot? Would it be possible to provide training about this matter for members of the institution (including members of the ethics committee)?

In these cases, it is not possible to provide a solution that would be suitable to all research and innovation institutions, but each institution should have a bearer of responsibility that fits its ethics infrastructure. Additionally, for successful implementation, a transition period is recommended to carry out communication activities and to enable the institution members to take part in the educational activities.

To sum up and return to the action and implementation plans, a question that needs further elaboration -- and not just being marked in proposed plans -- is whether the writers of the plan are in a position to write down the specific bearers of responsibility for each task and area or whether the bearers of responsibilities have to be specified by leaders of the institution at a later point. Additionally, it should be realized that there might be some issues that can only be dealt with to some extent at the institutional level as they may also require being addressed at the national level (e. g. national policy). Limitation of consideration to only one level may hinder the impact of creating and maintaining a culture of ethics and integrity.

6 Conclusion

Ethics infrastructure and governance are needed to safeguard trust in science. Ethics committees have an important role in the ethics infrastructure ensuring ethical research and innovation. The infrastructure must function at the institutional level and there should

be cooperation with the national level. The tasks of the ethics committee(s) within the research and innovation institution can vary, from issuing approvals to investigating (possible) misconduct to governance of the system. An example of an ethical system of governance for research and innovation in organizations developed was introduced based on the results of the H2020 ETNHA project, with a special focus on activities related to ethics committees. Additionally, for successful implementation of the system of ethical governance in a research and innovation institution, an action plan should be followed. Through involvement, and communication of policies and procedures, the plan should safeguard that both researchers and the public have trust in ethics infrastructure and ethics committees.

Acknowledgments. I would like to thank the reviewers of this chapter for their helpful suggestions and Tiina Kirss for her help with English expression.

References

1. Singapore Statement (2010). https://wcrif.org/guidance/singapore-statement
2. Forsberg, E.-M., et al.: Working with research integrity—guidance for research performing organisations: the Bonn PRINTEGER statement. Sci. Eng. Ethics **24**(4), 1023–1034 (2018). https://doi.org/10.1007/s11948-018-0034-4
3. Drenth, P., et. al.: Fostering research integrity in Europe: a report by the ESF Member Organization Forum on Research Integrity (2010). https://www.esf.org/fileadmin/user_upload/esf/ResearchIntegrity_Report2011.pdf
4. Hermeren, G., et al.: Mutual Learning Exercise (MLE) on Research Integrity. Processes and Structures (2019). https://ec.europa.eu/research-and-innovation/sites/default/files/rio/report/MLE%2520RI_Final%2520Report_0.pdf
5. Davies, S.R.: An ethics of the system: talking to scientists about research integrity. Sci. Eng. Ethics **25**(4), 1235–1253 (2018). https://doi.org/10.1007/s11948-018-0064-y
6. Six, F., Lawton, A.: Towards a theory of integrity systems: a configurational approach. Int. Rev. Adm. Sci. **79**, 639–658 (2013). https://doi.org/10.1177/0020852313501124
7. Anderson, M.S.: Global research integrity about the United States: research integrity infrastructure. Account. Res. **21**, 1–8 (2014). https://doi.org/10.1080/08989621.2013.822262
8. Foeger, N., Zimmerman, S.: Research integrity: perspectives from Austria and Canada. In: Bretag, T. (ed.) Handbook of Academic Integrity, pp. 809–821. Springer, Singapore (2016). https://doi.org/10.1007/978-981-287-098-8_41
9. Faintuch, J., Faintuch, S.: Past and current status of scientific, academic, and research fraud. In Faintuch, J., Faintuch, S. (eds.) Integrity of Scientific Research, pp. 3–8, Springer, Cham (2022). https://doi.org/10.1007/978-3-030-99680-2_1
10. Ngai, S., et al.: Haunted manuscripts: ghost authorship in the medical literature. Account. Res. **12**, 103–114 (2005). https://doi.org/10.1080/08989620590957175
11. Shaw, D.M.: The ethics committee as ghost author. J. Med. Ethics **37**, 706 (2011). https://doi.org/10.1136/medethics-2011-100120
12. Ding, D., et al.: Duplicate and salami publication: a prevalence study of journal policies. Int. J. Epidemiol. **49**, 281–288 (2020). https://doi.org/10.1093/ije/dyz187
13. Iphofen, R.: Regulating research. In: Iphofen, R. (ed.) Handbook of Research Ethics and Scientific Integrity, pp. 17–32. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-16759-2_52

14. Braun, R., Ravn, T., Frankus, E.: What constitutes expertise in research ethics and integrity? Res. Ethics **16**, 1–16 (2020). https://doi.org/10.1177/1747016119898402
15. Emanuel, E., Wendler, D., Grady, C.: An ethical framework for biomedical research. In: Emanuel, E.J., et al. (eds.) The Oxford Textbook of Clinical Research Ethics, pp. 123–135. Oxford University Press, Oxford (2008)
16. WMA (World Medical Association): Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects (1964/2013). https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/
17. Schrag, Z.M.: Ethical imperialism. Institutional Review Boards and the Social Sciences, 1965–2009. Johns Hopkins University Press, Baltimore (2010)
18. Israel, M.: Research Ethics and Integrity for Social Scientists, 2nd edn. SAGE, London (2015)
19. Council of Europe: Guide for Research Ethics Committee Members (2012). https://www.coe.int/t/dg3/healthbioethic/activities/02_biomedical_research_en/Guide/Guide_EN.pdf
20. Iphofen, R.: Ethical Decision-Making in Social Research. A Practical Guide. Palgrave Macmillan, London (2011)
21. Resnik, D.B.: The Ethics of Research with Human Subjects. Protecting People, Advancing Science, Promoting Trust. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-68756-8
22. Department of Health and Human Services: Institutional Review Boards. A Time for Reform (1998). https://oig.hhs.gov/oei/reports/oei-01-97-00193.pdf
23. Lynch, H.F., et al.: "We measure what we can measure": struggles in defining end evaluating institutional review board quality. Soc. Sci. Med. **292**, 114614 (2022). https://doi.org/10.1016/j.socscimed.2021.114614
24. Scherzinger, G., Bobbert, M.: Evaluation of research ethics committees: criteria for the ethical quality of the review process. Account. Res. **24**, 152–176 (2017). https://doi.org/10.1080/08989621.2016.1273778
25. Trace, S., Kolstoe, S.E.: Measuring inconsistency in research ethics committee review. BMC Med. Ethics **18**, 65 (2017). https://doi.org/10.1186/s12910-017-0224-7
26. Carniel, J., et al.: The ethics review and the humanities and social sciences: disciplinary distinctions in ethics review processes. Res. Ethics. (2022). https://doi.org/10.1177/17470161221147202
27. Dawson, A., et al.: Why research ethics should add retrospective review. BMC Med. Ethics **20**, 68 (2019). https://doi.org/10.1186/s12910-019-0399-1
28. Gonzales-Esteban, E., et al.: D.4.2. Final Concept of the ETHNA System (2022). https://ethnasystem.eu/wp-content/uploads/2022/02/D4.2_ETHNA_final-concept.pdf
29. National Bioethics Advisory Commission: Ethical and Policy Issues in Research Involving Human Participants: Recommendations of the National Bioethics Advisoty Commission. IRB: Ethics; Hum. Res. **23**, 14–16 (2001). https://doi.org/10.2307/3564051
30. ENRIO handbook: Recommendations for the Investigation of Research Misconduct (2019). http://www.enrio.eu/wp-content/uploads/2019/03/INV-Handbook_ENRIO_web_final.pdf
31. University of Liverpool: Research Integrity and Governance Committee (2023). https://www.liverpool.ac.uk/governance/university-committees/research-integrity-and-governance-committee/
32. Zwart, H., ter Meulen, R.: Addressing research integrity challenges: from penalising individual perpetrators to fostering research ecosystem quality care. Life Sci. Soc. Policy **15**(1), 1–5 (2019). https://doi.org/10.1186/s40504-019-0093-6
33. Horbach, S.P.J.M, Sorensen, M.P.: How to create and implement a Research Integrity Implementation Plan (RIPP). A guideline (ver. 2.0) (2022). https://sops4ri.eu/wp-content/uploads/Implementation-Guideline_FINAL.pdf

34. SOPs4RI: Guideline for Promoting Research Integrity in Research Funding Organisations (2022). https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/guideline-for-promoting-research-integrity-in-research-performing-organisations_horizon_en.pdf
35. SOPs4RI: Template for writing a research integrity promotion plan for research performing organisations (2022). https://sops4ri.eu/wp-content/uploads/Template-Research-Integrity-Promotion-Plan-RPOs_FINAL.pdf
36. EU Grants: How to complete your ethics self-assessment (2021). https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment_en.pdf
37. Stokel-Walker, C.: ChatGPT listed as author on research papers: many scientists disapprove. Nature **613**, 620–621 (2023). https://doi.org/10.1038/d41586-023-00107-z

Responsible Research and Innovation Learning Facilitation

Elisa Garcia-Garcia(✉) and Xavier Eekhout

Fundación Española para la Ciencia y la Tecnología, Pintor Murillo, 15, 28100 Alcobendas, Madrid, Spain
elisa.garcia@fecyt.es

Abstract. This chapter aims at supporting the institutionalization of Responsible Research and Innovation (RRI) principles by focusing in the facilitation of learning to research staff. We highlight the need to identify the learning objectives that want to be achieved by taking into account how researchers shape their ethical perspectives, as many of their behaviours are learned through informal training setups (e.g., supervisors, peers, etc.). This is complemented by reviewing the different approaches that can be applied to achieve the learning goals, and by showcasing existing public resources that can be further explored and tailored to plan specific interventions depending on the nature and needs of each institution.

Keywords: RRI · Research Integrity · Learning · Capacity Building · Researchers

1 Introduction

The ETHNA System aims at ensuring that research and innovation activities are carried out in a responsible way. For its implementation, and thus achieving the institutionalisation of Responsible Research and Innovation (RRI) principles, the system acknowledges two dimensions that need to be addressed, namely leadership -i.e., the top-down support- and base -i.e., the organisation's research staff with their values awareness, skills, knowledge, and practices. Regarding the latter, although the training of research staff was not directly addressed throughout the implementation of the ETHNA project, the current chapter builds from the perspective of a researcher career development support service to provide some recommendations for planning learning activities that foster ethical behaviour among research staff, aiming at strengthening the base of an institution wanting to institutionalise RRI. The overall considerations to have in mind when planning the learning activities are complemented with references leading to specific resources and examples.

Two distinct, interconnected motivations have propelled the use of structured learning of ethical behaviour of researchers: On the one hand, the rise of misconduct cases and frauded research, data irreproducibility and/or inaccessibility, the discrimination cases against women and minorities, or the unethical behaviour due to unbalanced power distribution within the research system —among other issues— have promoted the surge

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 137–154, 2023.
https://doi.org/10.1007/978-3-031-33177-0_9

of more structured courses and activities for promoting ethical conducts [1]. On the other hand, the change of paradigm of the connection of science and society, in which science not only serves society but is actually immersed in society. This way of understanding science has also prompted activities and interventions to increase the sense of responsibility of researchers toward society, including the need to embed society in the discussions and research decision-making [2]. These motivations have, for example crystallised in the new European Code of Conduct for Research Integrity, which includes clear calls to the need of having individuals trained to perform ethically [3].

Traditionally, knowledge values and ethical practices shaping the research conduct have been occurring through practice, through on-the-job learning processes. Thus, the promotion of ethical behaviours from researchers towards their community and towards society many times is based on implicit and observational learning. Nevertheless research has proved that ethical approaches toward research can also be learnt in more explicit ways [4]. More importantly, the values and moral judgement of Science in Society can also be stirred in structured processes that facilitate learning on the socio-ethical dimension of the research activity [5, 6].

This implies that learning RRI goes beyond the comprehension of established codes of conduct of a given organisation. It also implies learning the individual and community meaning-making of the established and agreed norms, the acquisition of the competencies needed to behave within these norms, and the motivation of individuals to adapt their behaviours to comply with the ethical common shared values of that organisation, the wider research community and society in general.

Moreover, research and innovation are by definition at the forefront of providing new knowledge, new processes, and new products. This way, it becomes of key importance to train individuals to perform ethically when encountering the unknown by taking into account not only individual and institutional values, but also the values shared with society. And thus, co-creating or collaboratively re-shaping codes to adopt the novel possible outcomes within society [7].

Due to the nature of RRI, learning should be facilitated to any personnel connected to research and development (R&D) processes in the organisations: researchers, research support staff, research managers, and personnel in contact with a sensitive issue of RRI. Furthermore, it is important to recall that RRI considers society as an active performer of science; hence, also society should be embedded in this learning process [2, 7]. Whenever individuals are involved in any aspect of science, they should receive training on, at least, some aspects of RRI. This aspect is partially covered in Science Education, a key area of RRI that this book does not articulate. In this chapter, we will mainly address researchers as the target of the learning process. Our aim with this chapter is providing guidance on how to design formal activities that render ethical learning and enhance the practices of RRI among researchers. From the functional areas of RRI (Research Integrity, Open Access and Open Data, Gender, Science Education, and Science Engagement) we have mainly focused on the Research Integrity (RI), a dimension that is further addressed in this book (see chapter by Marušić, Chap. 10, in this volume). Extensive literature and guidelines have already been created concerning Gender and Open Access and Open Data [8–10].

We have compiled existing literature, giving space to different perspectives of ethical learning instruction. Instructional systems design methodology reflects on what researchers need to learn to promote effective learning. However, particularly in the case of RI, only some factors are covered by formal learning. As mentioned previously, unstructured informal learning on-the-job occurs within the research praxis, where peers or senior researchers are a source of ethical information and education [11, 12]. This way, observational learning, i.e., role modelling, is also an important source of learning that needs to be considered. For this reason, to encourage the ethical behaviour of researchers, the community and organisation's ethical behaviour needs to be shaped and training must not be a standing-alone tool to promote RI within the research community; it should be considered part of the strategy.

Finally, we strongly encourage subscribing all training to the umbrella office or transversal management entity within the institution that deals with the aspect of RRI or to the RRI office(r) as proposed by the ETHNA system. The same would apply in the case of other RRI dimensions, for example, training or activities on gender should be linked to the gender office or related unit that deals with these issues. Notwithstanding this capitalisation of resources, these umbrellas are standpoints to address and seek support for self-organised activities, avoiding dispersion and parallelism and increasing visibility and efficiency.

2 Identifying the Learning Objectives

Instructional theory aims to describe how to help people learn and develop, creating conditions that boost the chance of learning and improve instruction [13]. This implies the need of understanding the purpose of the training, that will manifest in learning objectives, for designing the instructional process and assessments. In the case of RI, the instructor needs to analyse the learner's needs and aptitudes at the beginning of the design, and the knowledge, skills, and attributes that the learner need to incorporate into their ethical mental models to behave ethically during the research praxis [14].

Concerning the needs of the learners, it is advisable to be aware that the researchers within a given organisation will vary among other things in career stage and professional experience, for example, early-career researchers may not need to be proficient in processes involving conflict of interest when evaluating grant peers or research [12]. Learner's needs are also shaped by the discipline, for example, agri-food researchers need to be aware of ethical issues, regarding Genetic Modified Organisms (GMO). Related to learner's aptitude, assessing prior knowledge and tailoring training to various groups is a strategy to increase the learning effectiveness [15].

The discipline, career level, or key areas of RI will shape many aspects of the learning objectives. However, transversal learning objectives in RI training can be extracted from available literature and be outlined as follow:

- Acquisition of explicit knowledge of the organisation's values, enacted in codes of conduct, good practice lists, and various existing regulations and procedures that ensure ethical behaviour and responsible research.
- Internalisation and integration of the above-enacted regulations of the core values into the organisation's daily praxis.

- Acquiring a mindset ready to identify grey areas or questionable behaviour that can occur during their daily activity and can collide with the existing code of conduct or core values. In other words, providing the skills to recognize the ethical aspects of their everyday praxis and increase their decision-making skills.
- Increasing the responsive capacity of researchers to the broader societal dimension of their work, or in other words, fostering awareness about the fact that responsible research is a matter of the community and society.
- Increasing sensitivity to alternative perspectives and diversity of thinking, promoting cognitive complexity, cultural diversity, and empathy.
- Understanding cognitive bias, detecting and preventing moral justification and unethical decision-making.
- Incorporating intangibles as shared values of a team. These intangibles can come from new process or products (i.e., novel outcomes from the research and innovation activity) and/or from conflicting new approaches to research.
- Acquiring skills that enable the researchers to comply with the ethical conduct of research and responsible research and the organisation's shared values. For example, data management skills to comply with open data practices.

These general learning objectives are also aligned with the five processes which occur according to the RRI framework: anticipation, reflexivity, openness and inclusiveness, and adaptation [7], but discerning the specific skills, attributes, and knowledge that researchers should acquire to incorporate these processes is complex. It requires the understanding of the specific tasks that researchers need to perform ethically [13, 16].

We have distinguished three subsets of tasks, or processes, in the literature, depending on the approach to explain how RI is acquired, performed, or enhanced:

- Processes to achieve a social-ethical integration of research.
- Processes to comply with codes of conduct or regulations that secure ethical behaviour during the research praxis, and, finally,
- processes to take decisions ethically and avoiding biases.

Next, we will describe these approaches towards RI as a vehicle to highlight some of the competencies that can be the aim of training. In other words, we will expose the processes to perform RI, depending on the approach, that will frame the learning objectives.

2.1 Processes to Achieve a Social-Ethical Integration of Research

This approach to RI stem from an ethicists and philosophers' perspective. It aims to embed the ethical component from the beginning, during their praxis, or within the forecasting of the impact of research results on society. In this approach, researchers are expected to acquire moral awareness of the research activity first towards their own community, and second towards society as a whole [2].

In an instructional method based on this approach, the midstream modulation, authors describe that this trans-dimensional moral awareness is acquired when researchers engage in two types of reflection: first-order and second-order of reflection [17, 18]

According to the authors, in the first order of reflexivity, individuals reflect on their daily activities and interactions with peers, microethics, and ethical issues within their embedded research system. In the second order of reflexivity, individuals reflect on the impact of their decision and research activity on society.

Running a living lab -a methodology proposed for the implementation of the ETHNA system itself- is also an approach to RI that highlights the importance of the social-ethical integration of research. This methodology is based on the iterative feedback process with different stakeholders to foster ethical research and innovation (R&I) organisations. Key competencies, or clusters of competencies, to reinforce through training to support the tasks needed to run a living lab include design thinking skills [19], science communication, and public engagement skills; all of which are needed when involving other stakeholders of society in the promotion of reflection activities [20–22].

2.2 Processes to Comply with Codes of Conduct or Regulations

As mentioned above, many facts have propelled the redaction of codes of conduct, guidelines, or good practice to frame the standards of research conduct, for example because certain aspects of ethical researcher praxis are directly affected by laws, i.e., General Data Protection Regulation (GDPR) or animal welfare. The ETHNA system proposes Code of Ethics and Good Practices in R&I as one of the guidance tools for its institutional implementation, (see chapter González-Esteban, Chap. 2). Courses teaching these predefined rules to researchers are designed under two precepts: deficit of information can lead to misconduct, and compliance will guide a moral decision [23]. A salient task to comply with codes and regulations, and hence, research ethically according to this approach, is knowledge acquisition [24].

Although this transfer of information of existing rules can increase moral awareness, data show that this training impacts knowledge acquisition, but not moral judgement or ethical decision-making [4, 25, 26] processes that correlate higher with ethical behaviour. This fact can be explained by the phenomenon of reactance, a negative reaction towards externally imposed knowledge [27] and by the fact that researchers may feel that there is no answer in these codes to many of their experienced ethical dilemmas, so attending these courses could become a “ticking-off the-course-box” reaction [28–30]. Even with these dangers in mind, compliance courses on existing codes should not be discouraged, but should be combined with activities that aim at increasing decision-making skills, which will allow to better frame the acquisition of skills to knowledge based on shared previously agreed standards [31].

2.3 Processes to Take Decisions Ethically and Avoiding Biases

Although the knowledge of guidelines (e.g., through compliance courses, as mentioned above) and the awareness on topics subject of ethical issues affecting RI (e.g., doing research with other colleges, ethical leadership, supervisor/trainee relationship -see Sect. 4 for more examples), research looking into the ethical behaviour context at research institutions has identified that researchers -especially at early stages- face situations at the boundary of these rules. In these cases, the identification of the situation as an ethical

matter may not be a straightforward exercise [23, 32]. Considering these ambiguous contexts, ethical decision-making becomes a key process to research ethically and hence, the focus of large amount of research and training [24, 33]. By moral or ethical decision-making, we refer to decisions, including judgments, evaluations, and response choices that are related to moral issues [34].

As stated above, to propose effective learning objectives we need to understand the subjacent skills, knowledge, and attitudes of higher-order decision-making processes. Multiple and diverse theories have been proposed to explain how ethical decision-making works, develops and is connected to ethical behaviours. The main perspectives have been cognitive-development theoretical approaches, including moral reasoning, [35, 36] and affective theories [37], social intuitionist theories [38, 39], and social neuroscience theories, [40, 41]. A detailed exposition of this approaches exceeds the purpose of this book. However, more recently, a framework has been proposed to integrate many aspects of these theories into the social information processing theory [41] that it is worth mentioning in this chapter (Fig. 1). This framework contemplates the cognitive, affective, and social components that can modulate ethical decision-making, providing a useful reference to design training on RRI.

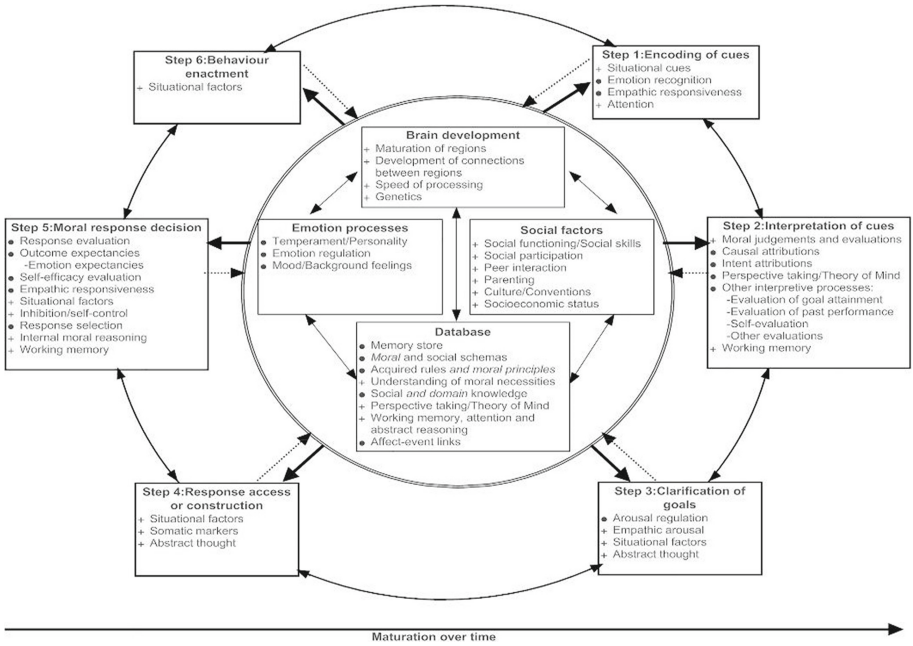


Fig. 1. The Social Information Processing-Moral Decision-Making Framework (SIP-MDM) from Garrigan et al. 2018

Training aiming to facilitate the learning of the components in this framework have an impact in the decision-making skills [32]. An example of a training activity considering many of the components of this framework would be presenting a dilemma case

to learners; then the facilitator guides them to attend to emotional cues and situational contexts; guides them also to interpretate who is responsible, trying to understand intentions, and facilitate them to take perspective from the role of the stakeholders within the case; at the same time, allow them to extrapolate the case with their daily praxis, and to contrast this information with their moral and social schemes acquired by previous experience while reflecting on the emotion that this situation evokes on them.

Some trainings have more activities focus on learning facilitation of sensemaking skills, including integrating and interpreting factors on moral dilemmas [32], self-reflecting on past experiences, predicting potential outcomes for the self and stakeholders, perspective taking [42], and emotional regulation [43]. In these approaches, participants also learn strategies to tackle ethical problems [32, 44]. Other activities of training has emphasized the social skills component of moral decision-making [45], and others have stressed the affective elements on the affective moral component [46].

On the other hand, training focusing on theoretical knowledge of virtues and ethical principles is framed in the "reasoning" model of morality acquisition, where the learner lacks the knowledge of the concepts of morality and virtues, and once learnt, ethical behaviour will follow [36, 47]. This focus has proved being less effective [4, 48].

Tasks Aiming at Bias Awareness. Like in any other decision-making process, biases also play a role in ethical decision-making [49]. An individual will likely use heuristics[1] to interpret the situation upon exposure to ethical dilemmas. This type of information-processing shortcut can introduce errors or bias in moral making sense of the situation or in forecasting the implications and consequences of the possible ethical scenarios [40]. They have implications on ethical behaviour because if bias occurs, then, even good-intentioned individuals with high information processing skills, moral awareness, and pro-social skills can still be engaged in wrongdoing [50, 51]. Training addressing this aspect in ethical decision-making has proven effective among business [52], health care [53], and engineering professionals [54]. In the case of bias training for researchers, there are only a few studies with still no clear results [55].

In summary, when designing training in the domain of responsible research, different learning objectives can be derived depending on the different research perspectives on how individuals acquire ethical competencies. This fact will frame the design of the activities because different instructional methods will achieve specific learning objectives [4]. Most probably, a combination of aspects must be explored [56]. For example, in research integrity the combination of a compliance-approach and behavioural approach should be considered [31].

3 Designing the Learning Activities

As previously discussed, designing a given training activity will be profoundly influenced by the learning objectives that we want to achieve [13]. Learning aspects, such as how

[1] Any approach to problem solving or self-discovery that employs a practical method that is not guaranteed to be optimal, perfect, or rational, but is nevertheless sufficient for reaching an immediate, short-term goal or approximation.

individuals learn more efficiently, will also affect the way the activity will be designed and delivered. Literature revision shows that not only learning reasons drive instructional design of research integrity training, but also the capacity of the organisation, the skills of staff, and even, existing inertias and culture of the learners, teachers, and organisations [57].

In the following paragraphs, we will outline different teaching-learning approaches detected in RRI training literature.

3.1 The Learning Approach Towards Research Integrity

Two distinct approaches to the teaching-learning process can be distinguished and selected when designing the activities: a lecturing approach or a facilitation approach. Its the purpose of this chapter to avoid revisiting these approaches thoroughly and to prevent favouring one over the other, as often they occur concurrently by combining activities within the same training or are intermixed in the same activity [58]. We rather urge the instructor or designer's skills to be aware of the selected approach in terms of the learning objectives and the most effective pedagogical techniques useful for the learning process.

In a lecturing approach, the teacher aims to transmit knowledge to the audience. Factors that impact the effectiveness of lecturing are mainly the structure of the delivered information and the teacher's communication skills, as it follows a teacher-centred approach [58]. The outcomes of the lecture approach can improve when the delivered information is related to previous knowledge of the learners and is interconnected with their professional reality, creating a conceptual map through which the new knowledge is connected with practices in real-world tasks. Furthermore, lecturing approach tends to be more efficient when direct engagement of the learner is promoted [59]. These aspects are important to be considered when designing online lectures or e-learning materials, as face-to-face interaction with the teacher and peers is reduced, if not lost.

In a facilitation approach, the teacher becomes a facilitator of the learning process. The basis of facilitation is that learning is a goal-oriented dynamic process of mutually shared existing knowledge, and a co-creation of knowledge that is integrated through critical reflection of the learner with the facilitation of the teacher [60]. This approach has its basis in humanistic learning theories and is related to adult learning principles. In this respect, the learning is constructed upon an experience that participants critically analyse, from a cognitive and an emotional point of view, which ultimately leads to internalisation of knowledge through deep reflection. This way, facilitation follows a student-centred approach, where individual learning styles can take place, and which necessarily implies the learner's active participation [61]. To achieve effective learning through a facilitation approach, experiences or activities should also be contextualized and related to the learner's own experience. This also implies the need to consider constructing a safe learning environment so that mutual knowledge can be freely exchanged. Finally, reflection should be guided so that new knowledge can arise after the group's exchanges on their perspectives [62]. When this learning approach is selected, the instructor's skills and experience as a facilitator are key to achieving deep learning [13], as there are many different methodologies which can be applied to facilitate the dialogue around research integrity and ethics [63].

Design thinking [64] and learning approaches based in ethnographic methods, like midstream modulation [17] have also been applied to instructional training design, although less literature is available on RI. These two approaches are used to facilitate learning in competencies related to the co-creation of common values and understanding of socio-ethical aspects.

3.2 Training Activities and Delivery Method

The lecture itself is the main activity when a lecturing approach is selected for the training of the researchers. As stated before, a lecture can increase its effectiveness through several interventions/activities looking for the engagement of the learner. Some are tests or quizzes, questioning and recurrent feedback, and related real examples. E-learning material framed as a lecture can also benefit from these interventions [59].

A key principle that many instructional designers work with is the design of problem-based activities [65]. In RI training, the use of dilemmas has been the most common strategy to create a problem situation [66]. Depending on the competencies that want to be trained, the dilemma's presentation and the tasks the learner will be asked to perform will be different.

More specifically, a frame activity widely used in RI training is based on cases, which is a way to attempt getting closer to professional experience of the learner [67]. Cases are contextualised dilemmas that invite the learners to discuss (i.e., applying a dilemma relevant to a specific field of knowledge and/or research activity). The case represents the ill-defined ethical situations at the boundaries of standards, regulations, and codes of conduct that researchers are exposed to in their daily praxis. Sensemaking skills, forecasting, reflexivity, and emotional regulation are some of the skills that can be learnt and trained though a facilitation approach [42, 43]. The cases can be tailored to increase proximity to the learners, enhance the emotions, highlight the dilemma's social implications, or explore the dilemma's social contexts.

Case-based activities vary depending on the learning competencies and the factor affecting learning. In this way, innovative ways to increase the situational experiences, like role-playing [68, 69] or fishbowl techniques [70], practice the perspective-taking skills of the learners. Others enhance the emotional aspects of dilemmas, delivered through storytelling [71, 72] or professional actors [73]. Finally, other activities enhance the engagement of learners through gamification [74] or even through interactive narratives in the form of video games [75, 76] or interactive films [77]. See Table 1 for specific examples.

In case-based studies, a key element is the discussion that the dilemma triggered, which needs to be structured. Otherwise, the learning objectives cannot be efficiently reached [61]. One example of structured dialogue can be found in the card-based method RESPONSE_ABILITY [30]. In this game, reflection also explores the conflict between morals and practices and the attribution of responsibility to organisational aspects rather than solely individual behaviour.

Table 1. Selection of frequent used strategies to enhance case-based activities.

Strategies to enhance case-based activities	Aim of the enhancement	Examples
Role playing	Increasing situational experiences Practicing perspective-taking Practicing emotional intelligence	RCR Casebook: Stories about Researchers Worth Discussing [69]
Fishbowl and debate	Practicing perspective-taking	The Dilemma Game [74]
Storytelling	Increasing empathy Increasing emotional aspects Practicing sensemaking skills	Path2Integrity courses (series Y) [72]
Theatre play	Increasing empathy Increasing emotional aspects	ConScience App [73]
Interactive films	Engagement of participants	"The Lab: Avoiding Research Misconduct" [77]
Gamification	Engagement of participants	Wicked Problems [75]
Structured debriefing	Increasing structured reflection	RESPONSE_ABILITY [30]

4 Deciding the Content

Along the present chapter we have tried to provide an overview of the elements that can drive the ethical behaviour of researchers (i.e., social-ethical integration of research, knowledge on codes of conduct, or decision-making), insisting in the idea that when planning to facilitate the learning of ethical research and innovation among research staff, the key step for deciding on the specific approaches (e.g., lecturing, facilitating or a combination of both) and methodology to apply (e.g., discussing a dilemma through a fish bowl exercise) lays on linking the learning objectives to the needs of the learners.

The diversity of content for RRI training is enormous, given the variety of topics and processes considered relevant for RRI. The main driver to select the topic of training is the learner's need. Organizations often prescribe or offer courses based on the expectations on what their researchers need to perform RRI. Discipline also shapes the content [24], focusing on specific sensible issues specific to a certain area of knowledge or R&I activity, e.g., GPDR when working with individuals, animal welfare when working with animals, or lab safety when handling chemicals.

Nevertheless, in order to support the reflection of the possible contents to be addressed through teaching-learning interventions, we show below research-related topics building upon the functional areas or dimensions of RRI (Fig. 2).

This is complemented with some examples of modules within courses, activities, or other setups dealing with research integrity; all of which could be adapted to address one or several specific RRI topics [33, 63, 78, 79].

Last but not least, along the recent years quite a number of projects and initiatives have worked in RRI, research integrity and ethics in R&I, including repositories of resources that can be used for implementing teaching-learning activities. With no intention of being exhaustive, we highlight several of them which can be very useful to anybody planning to implement institutional actions to foster ethical research and innovation behaviours among their research staff:

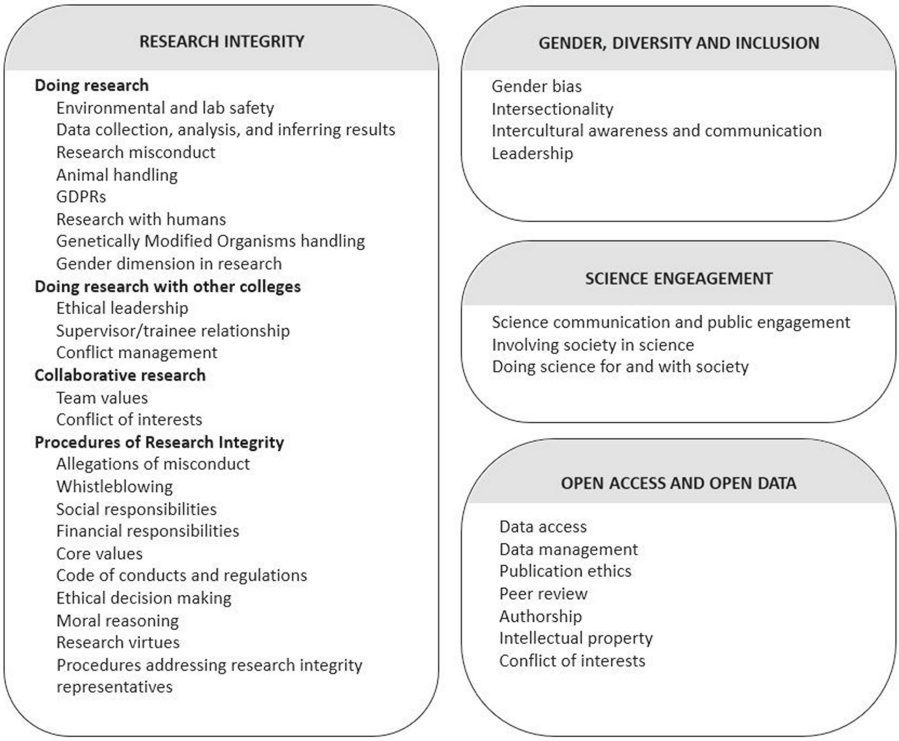


Fig. 2. Topics to be consider as content for RRI training, sorted by RRI area.

RRI tools started as an EU funded project under the 7th Framework Programme but is still active. It was proposed as a collaborative and inclusive project, with the aim of increasing creativity and shared ownership of the process, leading to an extensive repository of documents, including many training resources, particularly for the overall understanding of RRI [80].

SOPs4RI (Standard Operating Procedures for Research Integrity) has been another EU funded project, in this case under Horizon 2020 (H2020), with the aim of producing an online, freely accessible and easy-to-use 'toolbox' that can help research organisations to cultivate research integrity and reduce detrimental practice. A particularly interesting feature of this toolbox is the classification of training resources to different career stages and profiles [81].

Path2Integrity is another H2020 funded project supporting formal and informal learning methods contributing to establish a culture of RI, including the creation of units for learning RI. In this case, we would like to highlight their Research2Intergrity Roadmap [82], which not only compiles key references, but also the content that can be addressed and existing media to apply during teaching-learning exercises.

The Embassy of Good Science is a community platform developed offering support in handling day-to-day research practices and dilemmas. It contains guides, materials and a community that support training on research integrity and ethics [83].

5 Evaluation

After the delivery of the training is important to evaluate its effectiveness, not only for accountability but also to improve research findings on RI training and move the field forward [25]. The assessment design should take place at the beginning of the design cycle and not after [13]. Indeed, it is a consumable and difficult task of instructional designs. According to Kirkpatrick [84] training evaluation should assess the learners' reactions, achievements of the learning objectives, individual behavioural change, and organizational outcomes.

Reactions measure the level of satisfaction of participants after the training. Training organizations use that feedback to evaluate the effectiveness of the training, learners' perceptions, potential future improvements, and justification for the training expense. Normally, the kind of questions to answer would need to be tailored to the topic addressed in the intervention, but have structures such as:

- Did the trainees feel that the training was worth their time?
- Did they think that it was successful?
- What were the biggest strengths of the training, and the biggest weaknesses?
- Did they like the venue and presentation style?
- Did the training session accommodate their personal learning styles?

Nevertheless, general satisfaction is not a reliable measure of learning because it depends heavily on motivational aspects and expectations of the learners [85]. On the contrary, when the reactions related to a specific parts of the training (e.g., length of an specific activity, clarity of the content) then, the reaction response can be useful to tailor this specific sections.

Learning the objectives is really the key aspect to measure the impact on learning and will consist of evaluating knowledge acquisition. This requires the design of more sophisticated tests focused on whatever wants to be evaluated: knowledge (e.g., on a given code or protocol), skills (e.g., communication skills) and attitudes (e.g., leadership or ethical decision-making). These tests should be administrated before and after the training to be able to compare effects and the use of interviews or verbal assessments could also be applied. Furthermore, this "before and after" approach requires including control of population [14, 32].

Behavioural changes are difficult to measure in general, and even more in ethics-related trainings due to the low frequency of uncovered misconduct cases [84].

Organisations impact seeks to determine the tangible results of the learning intervention looking at things like reduced costs, higher efficiency, employee satisfaction, etc., so taking into account the heterogeneous and interconnected aspects of RI, matching organisational changes to specific training interventions seems unrealistic [84].

6 Overall Remarks

The disconnection between research and society cannot be a model of research praxis to follow. Not only is accountability to tax-payers a factor in considering a more fluent communication between research endeavours and society, but trust in research is also fundamental to create a knowledge-based society. In this respect, research and society are at stake if research does not follow ethical norms.

Several measures have been deployed in the R&I environment to ensure the ethical behaviour of researchers. The ETHNA project has proposed a flexible ethical governance system building upon several of these measures as guiding tools to achieve the institutionalisation of RRI principles (Code of Ethics and Good Practices in R&I, Ethics Committee on R&I and Ethics Line) together with a process (based in living lab methodology) and indicators to measure progress.

In this chapter we attempt to complement this approach to the governance system itself by providing recommendations on how to approach the learning of responsible research and innovation among the institution's research staff. In this sense, teaching-learning activities around RI would be considered as yet another of these measures.

Although expectations are high on explicit training actions (e.g., lectures, interactive workshops or compliance exercises), we cannot obviate that other types of learning are not explicit and occur by observing peers and senior researchers [11, 12]. It has been found that the perception of a healthy research environment has been linked with more desirable research practices [67], so cultivating a RI organisation environment is crucial to shaping the ethical behaviours of individuals [68]. Interviews and surveys have revealed that ethical behaviour has a high environmental component and does not rely only on individual capabilities. Increased stress, the competitiveness of the research funding models and of professional development, as well as power relationships, can shape the ethical behaviour of researchers [69]. Higher rearrangements need to be pursued to change the environment in which research occurs.

Due to this, along the current chapter we have attempted to provide an overview of the elements which influence the ethical behaviour of researchers, in order to provide a framework which should guide the planning of any teaching activities aimed at facilitating the learning of responsible research and innovation. The fact that RRI as a concept covers many areas and topics, as well as the need to take into consideration the nature of the target group addressed (e.g., institutional mission, research fields, activities implemented, level of experience, etc.), makes it impossible to propose a one-size-fits-all training module. Nevertheless, we expect that taking into consideration the possible learning objectives to address, the available approaches, plus the contents to be covered —together with the extensive resources already available, many of which are referenced here— this chapter will be useful for any person working in career development support of researchers, to plan specific interventions within their organizations.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission. The authors would like to thank the reviewers for their comments.

References

1. Godecharle, S., Fieuws, S., Nemery, B., Dierickx, K.: Scientists still behaving badly? A survey within industry and universities. Sci. Eng. Ethics **24**(6), 1697–1717 (2017). https://doi.org/10.1007/s11948-017-9957-4
2. Macnaghten, P.: Models of science policy: from the linear model to responsible research and innovation. In: Mieg, H.A. (eds.) The Responsibility of Science. SHPS, vol. 57, pp. 93–106. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-91597-1_5
3. The European Code of Conduct for Research Integrity. ALLEA - All European Academies, Berlin (2017)
4. Katsarov, J., Andorno, R., Krom, A., van den Hoven, M.: Effective strategies for research integrity training—a meta-analysis. Educ. Psychol. Rev. **34**, 935–955 (2022). https://doi.org/10.1007/s10648-021-09630-9
5. Tomić, V., Buljan, I., Marušić, A.: Development of consensus on essential virtues for ethics and research integrity training using a modified Delphi approach. Account Res., 1–24 (2022). https://doi.org/10.1080/08989621.2022.2128340
6. Besley, J.C., Dudo, A., Yuan, S.: Scientists' views about communication objectives. Public Underst. Sci. **27**, 708–730 (2018). https://doi.org/10.1177/0963662517728478
7. Stilgoe, J., Owen, R., Macnaghten, P.: Developing a framework for responsible innovation. Res. Policy **42**, 1568–1580 (2013). https://doi.org/10.1016/j.respol.2013.05.008
8. Bezjak, S., et al.: Open Science Training Handbook. (2018). https://doi.org/10.5281/ZENODO.1212496
9. UN Women Training Centre Annual Report: UN Women Training Centre, Santo Domingo (2016)
10. Lucy Fergurson: Feminist pedagogies in training for gender equality, Santo Domingo (2019)
11. Rissanen, M., Löfström, E.: Students' research ethics competences and the university as a learning environment. Int. J. Educ. Integr. **10**, 17–30 (2014)
12. Hyytinen, H., Löfström, E.: Reactively, proactively, implicitly, explicitly? Academics' pedagogical conceptions of how to promote research ethics and integrity. J. Acad. Ethics **15**(1), 23–41 (2016). https://doi.org/10.1007/s10805-016-9271-9
13. Goldstein, I.L., Ford, J.K.: Training in Organizations: Needs Assessment, Development, and Evaluation, 4th edn. Wadsworth/Thomson Learning, Belmont (2002)
14. Antes, A.L.: A systematic approach to instruction in research ethics. Account Res. **21**, 50–67 (2014). https://doi.org/10.1080/08989621.2013.822269
15. Hofmann, B., Bredahl Jensen, L., Eriksen, M.B., Helgesson, G., Juth, N., Holm, S.: Research integrity among PhD students at the faculty of medicine: a comparison of three Scandinavian universities. J. Empir. Res. Hum. Res. Ethics **15**, 320–329 (2020). https://doi.org/10.1177/1556264620929230
16. Thiagarajan, S., et al.: Instructional Development for Training Teachers of Exceptional Children: A Sourcebook. Indiana Univ., Bloomington. Center for Innovation in Teaching the Handicapped. National Center for Improvement of Educational Systems (DHEW/OE), Washington

17. Schuurbiers, D.: What happens in the lab: applying midstream modulation to enhance critical reflection in the laboratory. Sci. Eng. Ethics **17**, 769–788 (2011). https://doi.org/10.1007/s11948-011-9317-8
18. van de Poel, I., Zwart, S.D.: Reflective equilibrium in R & D networks. Sci. Technol. Human. Values **35**, 174–199 (2010). https://doi.org/10.1177/0162243909340272
19. Mohamed, T.I.: Refine the students' skills through design education to create ethical products. SAGE Open **11**, 2158244021998360 (2021). https://doi.org/10.1177/2158244021998360
20. Johnson, A.R., Nadkarni, N.M., Weber, C.Q.: Providing public engagement training to build connections between the community and research ethics professionals: a pilot project. J. Empir. Res. Hum. Res. Ethics **17**, 545–553 (2022). https://doi.org/10.1177/15562646221126282
21. Davies, S.R.: Constructing communication: talking to scientists about talking to the public. Sci. Commun. **29**, 413–434 (2008). https://doi.org/10.1177/1075547008316222
22. Weber, C., Allen, S., Nadkarni, N.: Scaling training to support scientists to engage with the public in non-traditional venues. J. Sci. Commun. **20**, 1–15 (2021). https://doi.org/10.22323/2.20040802
23. Sonenshein, S.: The Role of Construction, Intuition, and Justification in Responding to Ethical issues at Work: the Sensemaking-Intuition Model
24. Mulhearn, T.J., Steele, L.M., Watts, L.L., Medeiros, K.E., Mumford, M.D., Connelly, S.: Review of instructional approaches in ethics education. Sci. Eng. Ethics **23**, 883–912 (2017). https://doi.org/10.1007/s11948-016-9803-0
25. Steele, L.M., Mulhearn, T.J., Medeiros, K.E., Watts, L.L., Connelly, S., Mumford, M.D.: How do we know what works? A review and critique of current practices in ethics training evaluation. Account. Res. **23**, 319–350 (2016). https://doi.org/10.1080/08989621.2016.1186547
26. Watts, L.L., Medeiros, K.E., Mulhearn, T.J., Steele, L.M., Connelly, S., Mumford, M.D.: Are ethics training programs improving? A meta-analytic review of past and present ethics instruction in the sciences. Ethics Behav. **27**, 351–384 (2017). https://doi.org/10.1080/10508422.2016.1182025
27. Worchel, S., Brehm, J.W.: Direct and implied social restoration of freedom. J. Pers. Soc. Psychol. **18**, 294–304 (1971). https://doi.org/10.1037/h0031000
28. Phillips, T., Nestor, F., Beach, G., Heitman, E.: America COMPETES at 5 years: an analysis of research-intensive universities' RCR training plans. Sci. Eng. Ethics **24**(1), 227–249 (2017). https://doi.org/10.1007/s11948-017-9883-5
29. Sarauw, L.L., Degn, L., Ørberg, J.W.: Researcher development through doctoral training in research integrity. Int. J. Acad. Dev. **24**, 178–191 (2019). https://doi.org/10.1080/1360144X.2019.1595626
30. Felt, U., Frantz, F.: RESPONSE_ABILITY a card-based engagement method to support researchers' ability to respond to integrity issues. Sci. Eng. Ethics **28** (2022). https://doi.org/10.1007/s11948-022-00365-6
31. Mumford, M.D., Watts, L.L., Medeiros, K.E., Mulhearn, T.J., Steele, L.M., Connelly, S.: Biomedical ethics education may benefit from integrating compliance and analysis approaches (2016). https://doi.org/10.1038/ni.3465
32. Mumford, M.D., et al.: A sensemaking approach to ethics training for scientists: preliminary evidence of training effectiveness. Ethics Behav. **18**, 315–339 (2008). https://doi.org/10.1080/10508420802487815
33. Pizzolato, D., Abdi, S., Dierickx, K.: Collecting and characterizing existing and freely accessible research integrity educational resources. Account. Res. **27**, 195–211 (2020). https://doi.org/10.1080/08989621.2020.1736571
34. Turiel, E.: The Development of Social Knowledge: Morality and Convention. Cambridge University Press (1983)

35. Kohlberg, L.: The Philosophy of Moral Development Moral Stages and the Idea of Justice. Harper & Row (1981)
36. Rest, J.R.: Moral Development: Advances in Research and Theory. Praeger, New York (1986)
37. Hoffman, M.L.: Empathy and moral development: Implications for caring and justice. Cambridge University Press, New York (2000). https://doi.org/10.1017/CBO9780511805851
38. Kahneman, D.: Thinking, Fast and Slow. Farrar, Straus and Giroux (2011)
39. Hofmann, B., Holm, S.: Research integrity: environment, experience, or ethos? Res. Ethics **15**, 1–13 (2019). https://doi.org/10.1177/1747016119880844
40. MacDougall, A., Martin, A., Sahatjian, Z., Mumford, M.: A review of theory progression in ethical decision making literature. J. Organ. Psychol. **14**, 9–19 (2014)
41. Garrigan, B., Adlam, A.L.R., Langdon, P.E.: Moral decision-making and moral development: toward an integrative framework (2018). https://doi.org/10.1016/j.dr.2018.06.001
42. Harkrider, L.N., et al.: Improving case-based ethics training with codes of conduct and forecasting content. Ethics Behav. **22**, 258–280 (2012). https://doi.org/10.1080/10508422.2012.661311
43. Johnson, J.F., et al.: Case-based ethics education: the impact of cause complexity and outcome favorability on ethicality. J. Empir. Res. Hum. Res. Ethics **7**, 63–77 (2012). https://doi.org/10.1525/jer.2012.7.3.63
44. Kligyte, V., et al.: Application of a sensemaking approach to ethics training in the physical sciences and engineering. Sci. Eng. Ethics **14**, 251–278 (2008). https://doi.org/10.1007/s11948-007-9048-z
45. Kucukkelepce, G.E., Dinc, L., Elcin, M.: Effects of using standardized patients on nursing students' moral skills. Nurs. Ethics **27**, 1587–1602 (2020). https://doi.org/10.1177/0969733020935954
46. Morrill, D., Westrick, S.J.: Fostering nursing students' moral decision-making through use of an affective learning module. Nurse Educ. **47**, 236–240 (2022). https://doi.org/10.1097/NNE.0000000000001178
47. Tseng, H.H.: The development of the personal and professional values-integrated framework as an aid to ethical decision making. Am. J. Educ. Res. **3**, 1163–1167 (2015). https://doi.org/10.12691/education-3-9-16
48. Mumford, M.D., Steele, L., Watts, L.L.: Evaluating ethics education programs: a multilevel approach. Ethics Behav. **25**, 37–60 (2015). https://doi.org/10.1080/10508422.2014.917417
49. Medeiros, K.E., et al.: Biases in ethical decision making among university faculty. Account. Res. **21**, 218–240 (2014). https://doi.org/10.1080/08989621.2014.847670
50. Hirsch, P.B.: The serpent under it: cognitive bias in ethics and compliance training. J. Bus. Strateg. **43**, 406–409 (2022). https://doi.org/10.1108/JBS-08-2022-0145
51. Tomlin, K.A., Metzger, M.L., Bradley-Geist, J., Gonzalez-Padron, T.: Are students blind to their ethical blind spots? An exploration of why ethics education should focus on self-perception biases. J. Manag. Educ. **41**, 539–574 (2017). https://doi.org/10.1177/1052562917701500
52. Loe, T.W., Ferrell, L., Mansfield, P.: A review of empirical studies assessing ethical decision making in business. J. Bus. Ethics **25**, 185–204 (2000). https://doi.org/10.1023/A:1006083612239
53. Lahey, T.: Perspective: a proposed medical school curriculum to help students recognize and resolve ethical issues of global health outreach work. Acad. Med. **87**, 210–215 (2012). https://doi.org/10.1097/ACM.0b013e31823f3fb1
54. Finelli, C., et al.: An assessment of engineering students' curricular and co-curricular experiences and their ethical development. J. Eng. Educ. **101** (2012). https://doi.org/10.1002/j.2168-9830.2012.tb00058.x

55. Mecca, J.T., Medeiros, K.E., Giorgini, V., Gibson, C., Mumford, M.D., Connelly, S.: Biases and compensatory strategies: the efficacy of a training intervention. Ethics Behav. **26**, 128–143 (2016). https://doi.org/10.1080/10508422.2014.997878
56. Medeiros, K.E., Watts, L.L., Mulhearn, T.J., Steele, L.M., Mumford, M.D., Connelly, S.: What is working, what is not, and what we need to know: a meta-analytic review of business ethics instruction. J. Acad. Ethics **15**(3), 245–275 (2017). https://doi.org/10.1007/s10805-017-9281-2
57. Shephard, K., Trotman, T., Furnari, M., Löfström, E.: Teaching research integrity in higher education: policy and strategy. J. High. Educ. Policy Manag. **37**, 615–632 (2015). https://doi.org/10.1080/1360080X.2015.1102823
58. Jarvis, P.: The Theory and Practice of Teaching. Routledge (2006). https://doi.org/10.4324/9780203016442
59. Ambrose, S.A., Bridges, M.W., Dipietro, M., Lovett, M.C., Norman, M.K., Mayer, R.E.: 7 Research-Based Principles for Smart Teaching
60. Burrows, D.E.: Facilitation: a concept analysis. J. Adv. Nurs. **25**, 396–404 (1997). https://doi.org/10.1046/j.1365-2648.1997.1997025396.x
61. Regmi, K.: A Review of Teaching Methods -Lecturing and Facilitation in Higher Education (HE): A Summary of the Published Evidence (2012)
62. Țîru, C.M.: Styles of facilitation in the educational process at University Leve. J. Plus Educ. **XI**, 33–41 (2013)
63. Koterwas, A., Dwojak-Matras, A., Kalinowska, K.: Dialogical teaching of research integrity: an overview of selected methods. Facets **6**, 2138–2154 (2021). https://doi.org/10.1139/FACETS-2021-0045
64. Ribeiro, M.M., Hoover, E., Burford, G., Buchebner, J., Lindenthal, T.: Values as a bridge between sustainability and institutional assessment. Int. J. Sustain. High. Educ. **17**, 40–53 (2016). https://doi.org/10.1108/IJSHE-12-2014-0170
65. Merrill, M.D.: First principles of instruction. Educ. Technol. Res. Dev. **50**, 43–59 (2002). https://doi.org/10.1007/BF02505024
66. Kalichman, M., Brown, S.: Effects of training in the responsible conduct of research: a survey of graduate students in experimental sciences. Sci. Eng. Ethics **4**, 487–498 (1998). https://doi.org/10.1007/s11948-998-0041-y
67. Brock, M.E., Vert, A., Kligyte, V., Waples, E.P., Sevier, S.T., Mumford, M.D.: Mental models: an alternative evaluation of a sensemaking approach to ethics instruction. Sci. Eng. Ethics **14**, 449–472 (2008). https://doi.org/10.1007/s11948-008-9076-3
68. Brummel, B.J., Gunsalus, C.K., Anderson, K.L., Loui, M.C.: Development of role-play scenarios for teaching responsible conduct of research. Sci. Eng. Ethics **16**, 573–589 (2010). https://doi.org/10.1007/s11948-010-9221-7
69. RCR Casebook: Stories about Researchers Worth Discussing. https://ori.hhs.gov/rcr-casebook-stories-about-researchers-worth-discussing
70. Cummings, C.A.: Rethinking the fishbowl discussion strategy: a mechanism to construct meaning and foster critical thinking and communication skills though students' dialogue. J. Health Educ. Teach. Tech. **2**, 23–37 (2015)
71. Green, M., Brock, T.: The role of transportation in the persuasiveness of public narrative. J. Pers. Soc. Psychol. **79**, 701–721 (2000). https://doi.org/10.1037/0022-3514.79.5.701
72. Path2integrity Courses (Serie Y). https://learning-p2i.eu/course/index.php?categoryid=10
73. The ConScience App. http://www.hetacteursgenootschap.nl/productions/the-conscience-app/
74. Dilemma game. https://www.eur.nl/en/about-eur/policy-and-regulations/integrity/research-integrity/dilemma-game

75. Casas-Roma, J., Conesa, J., Caballé, S.: Teaching ethics in online environments: a prototype for interactive narrative approaches. In: Guralnick, D., Auer, M.E., Poce, A. (eds.) TLIC 2022. LNNS, vol. 581, p.p 496–509. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-21569-8_47
76. Nardo, A., Gaydos, M.: 'Wicked problems' as catalysts for learning in educational ethics games. Ethics Educ. **16**, 492–509 (2021). https://doi.org/10.1080/17449642.2021.1979283
77. The Office of Research Integrity, Interactive Movie on Research Misconduct. https://ori.hhs.gov/the-lab
78. Abdi, S., Pizzolato, D., Nemery, B., Dierickx, K.: Educating PhD students in research integrity in europe. Sci. Eng. Ethics **27**(1), 1–19 (2021). https://doi.org/10.1007/s11948-021-00290-0
79. Buljan, I., et al.: VIRT 2 UE: D2.1. Review of the existing ERI training literature and practices Project Details Project full name Virtue based ethics and Integrity of Research: Train-the-Trainer programme for Upholding principles and practices of the European Code of Conduct for Research Integrity List of reviewers (section 4) (2019)
80. RRI Tools training resources. https://rri-tools.eu/training/resources
81. SOPs4RI training resources. https://sops4ri.eu/tool_category/training/
82. Path2Integrity Roadmap. https://www.path2integrity.eu/teaching-RI
83. The Embassy of Good Science training resources. https://embassy.science/wiki/Training
84. Kirkpatrick, D.L.: Evaluating Training Programs: The Four Levels. Berrett-Koehler Publishers, San Francisco (1994)
85. Arthur, W., Tubre´, T., Paul, D.O.N.D., Edens, P.S.: Teaching effectiveness: the relationship between reaction and learning evaluation criteria. Educ. Psychol. **23**, 275–285 (2003). https://doi.org/10.1080/0144341032000060110

Activities to Foster Public Engagement in Research and Innovation. Examples from the NewHoRRIzon Project

Erich Griessler(✉) and Shauna Stack

Institute for Advanced Studies, Josefstädter Str. 39, 1080 Vienna, Austria
{Erich.Griessler,Shauna.Stack}@ihs.ac.at

Abstract. In this chapter we take results from the Horizon 2020 project, NewHoRRIzon to show that a variety of activities and approaches, addressing different levels and actors, are needed to spur public engagement in research and innovation. NewHoRRIzon spanned over all areas of the 8th European Framework Programme (Horizon 2020) and created 19 'Social Labs' to look into different research and innovation themes and their relation to the concept of Responsible Research and Innovation (RRI). This chapter highlights specific pilots stemming from this project which were designed by research and innovation stakeholders over the course of sequential labs, workshops, and meetings. The pilots featured here represent replicable activities and innovative ideas for researchers and research organizations to take up and use in their public engagement practices and policies.

Keywords: public engagement · Responsible Research and Innovation · RRI implementation · Social Labs · Stakeholder inclusion · Horizon 2020 · science in society

1 Introduction

Public engagement has been advocated for a very long time for democratic and scientific reasons [1–3]. Nevertheless, engaging with the public is still relatively uncommon in many disciplines [see for example, 4] and a disparity exists within Europe as to the quantity of public engagement activities [5]. Public engagement is still an ambiguous and multi-facetted buzzword [6] that carries different meanings for different stakeholders. It is based on different values [2] as well as concepts [7, 8], addresses different disciplines and takes many different forms that engage the public at different moments of research and innovation such as in science shops [9], events [10], consensus conferences [11], lay membership on scientific advisory committees, [12] experiments and demonstrations [13], to name a few.

In order to provide an overview on different formats of public engagement activities we present a number of activities that were developed in a bottom-up approach together with stakeholders from the research and innovation community. Thus, they are therefore

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 155–169, 2023.
https://doi.org/10.1007/978-3-031-33177-0_10

anchored in everyday needs, experiences, aspirations and institutional limitations of researchers and other stakeholders. In the NewHoRRIzon project, on which this chapter is based, we set out to implement the concept of Responsible Research and Innovation (RRI), which entails amongst other concepts public engagement, across all research funding programmes of the European Framework Programme H2020. For approximately two years, the project engaged with more than 720 stakeholders from research and innovation and developed, together with them, more than 50 pilot actions. Many of the pilot actions developed tried to foster public engagement as the participants in the project's 19 different so-called Social Labs felt the need to specifically address this issue. In this paper we will shortly explain the approach by which the public engagement pilots were created and will categorize and showcase some of them.

2 NewHoRRIzon and Social Labs

The NewHoRRIzon project tried to contribute to the implementation of the policy concept of Responsible Research and Innovation (RRI) within the 8th Framework Programme of the European Union, Horizon 2020 (H2020). Besides public engagement, the topic of this chapter, RRI entails other keys such as gender, ethics, science education and open access [14]. The RRI concept set out to implement these keys in research and innovation, thus making it more gender sensitive in research practice and topics, more considerate of ethical implications of, and conduct in research, and boosting science education, open scientific data, and results to the public.

The NewHoRRIzon project was an attempt to create, together with relevant stakeholders, measures and activities that would support RRI in research and innovation. To these aims, NewHoRRIzon modified the Social Lab concept developed by Zaid Hassan [15] as a process for solving complex societal problems with a bottom-up-approach of stakeholder engagement [16]. A Social Lab brings together diverse groups of stakeholders to focus on addressing complex societal challenges. The process involves a diverse group of stakeholders who are encouraged to contribute their unique perspectives to the challenge, the Social Lab Manager to own and manage the process and a facilitator to guide open ideation during a series of workshops [17, 18].

NewHoRRIzon started with mapping the stakeholders of H2020 and continued by analyzing the state of RRI in all Programme Lines of H2020 [19]. Thereafter, it created 19 Social Labs that covered all thematic Programme Lines of H2020.

The NewHoRRIzon partners developed a common manual that guided the Social Labs. The pilot action process for each Social Lab is documented in individual reports [20–23] and a Guide to Good Practice [24].

Altogether the Social Labs attracted more than 720 stakeholders from across Europe. The stakeholders came from research, research funding, civil society, policy making and business and in each of the Social Labs, spanning over two years, developed so-called pilots that addressed RRI challenges as the stakeholders perceived them in their own working environment.

The altogether 59 NewHoRRIzon pilots covered all RRI keys and addressed researchers, research funders, policy makers, representatives from business and civil society as well as citizens [25].

2.1 NewHoRRIzon Public Engagement Activities

Within the NewHoRRIzon project Social Labs came up with many ideas and pilots that focused on public engagement. Figure 1 below shows a map outlining these pilots based on their type of approach to Public Engagement. We categorized public engagement activities in (a) Exhibits about research, (b) Dialogical Formats, (c) Training Materials, (d) hands-on formats and (e) capacity building. A comprehensive description of the NewHoRRIzon pilot actions is available in an online brochure [25].

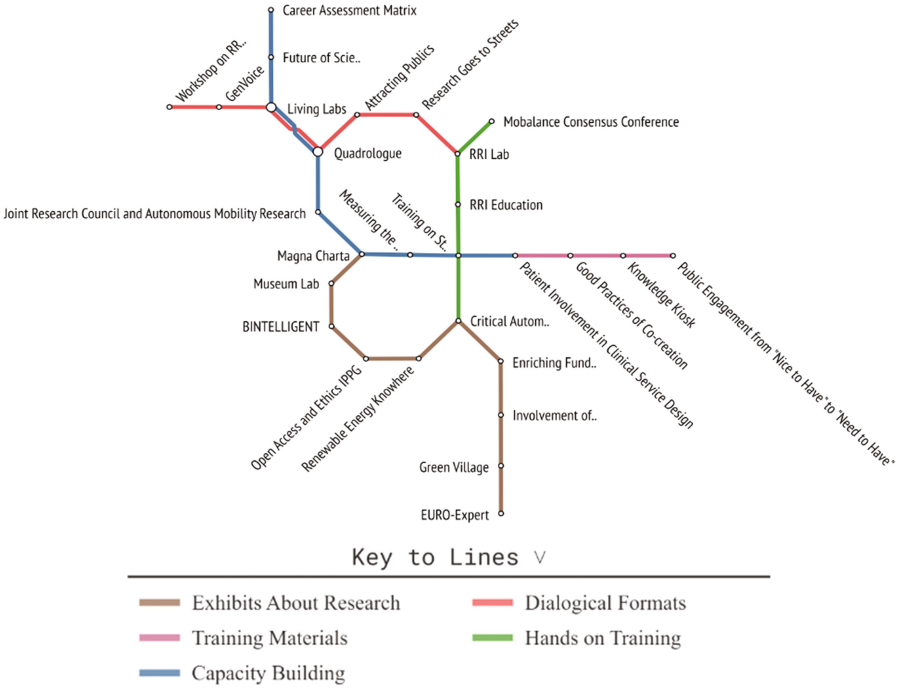


Fig. 1. Map of the different Public Engagement Pilot Actions developed in NewHoRRIzon grouped by different approaches. (Map designed with web tool: https://metrosets.ac.tuwien.ac.at/ [26])

As the map demonstrates, the approaches to public engagement overlap and diverge from each other in various ways. While some can be used in multiple settings, others were designed for more specific contexts and serve as inspiration for other organizations. A first group, Exhibits About Research, were meant to provide information about researchers' work and the work of their organizations to a broader public. Alternatively, Social Lab participants also created Dialogical Formats when they felt the need to engender and foster bidirectional exchange between researchers and the public/stakeholders. A third group of public engagement pilots concerned Training Materials that raise awareness about public engagement and enable researcher organizations to foster their own public engagement practices. Another category of pilots trained researchers in a 'hands

on' way in public engagement. The final group of activities addressed a perceived lack of organizational preconditions that hinder researchers from engaging with the public, an activity that is little credited in the assessment of organizations and individual careers.

In the following sections we will go into detail about some of the pilot actions shown in the map to illustrate the kinds of motivations, obstacles, lessons learned and transferrable practices that can be taken up from the NewHoRRIzon project.

2.2 Exhibits that Inform About Research

One category of pilots dedicated to public engagement includes exhibitions of information, or examples of researchers engaging with the public simply by sharing what they are doing in a way that considers the potential needs and uses of the stakeholders. In other words, these pilots represent an advanced level of research dissemination activity because their means of presentation are also meant to be practically relevant to stakeholders.

Euro-Expert and RRI. The pilot "Euro-Expert"[1] is a website and communicable output from legal scholarship and anthropology that works on making material accessible to people outside the scientific domain. Workshops during the NewHoRRIzon project showed that despite the relevance of legal research to civic life, researchers might lack expertise, time, or sense a general skepticism towards public engagement amongst their research community. This pilot tried to combat these obstacles to be more inclusive and adaptive to social changes within the subject and in the interface between researchers and legal practitioners.

The specific topic of the pilot is cultural expertise in legal theory and practice. In the development phase, it was important for the stakeholders to create a website to help engage people outside the scientific world, and specifically increase engagement between researchers and information users, such as legal professionals and people at court, in order to improve legal processes.

The website shares research results from legal and anthropological research about the role of cultural expertise in the context of legal decision-making and targets stakeholders such as cultural experts, judges, and prosecutors. These stakeholders as well as interested publics can easily get the latest relevant research results from a dedicated website and can even contribute insights through blogs. Such a pilot can be used as a model for other projects that want to engage with their wider stakeholder community to have their research taken up to improve the field of practice.

Renewable Energy Knowhere. The renewable energy field is constantly changing. Foundations, associations, small and big NGOs, organizations break up and suspend their operation, and more and more university departments and faculties take up the topic of sustainability and renewable energy. Accessibility of information is a crucial step when it comes to raising awareness of existing efforts in the field of renewable

[1] https://euro-resp.com (Downloaded 11.02.2023) [27].

energy and more ambitiously, to tackling the energy and climate crises. The Renewable Energy Knowhere pilot action helps stakeholders make sense of all this renewable energy research and institutional activity by summarizing EU countries' various renewable energy statuses.

At its core, the pilot is a database[2] in the form of a zoomable online-map (see Image 1) focusing on the Hungarian renewable energy field (Fig. 2).

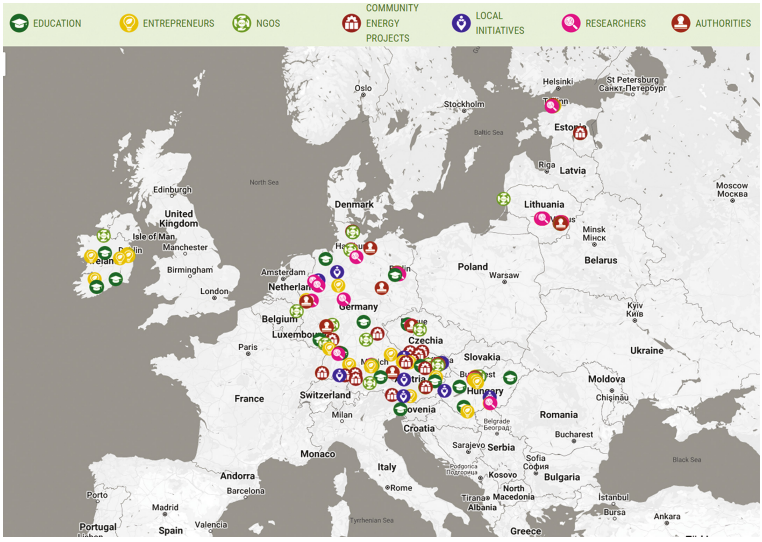


Fig. 2. Visual of map from the website "Renewable Energy Knowhere".

Included there are specific categories of interest such as education, entrepreneurs, NGOs, community projects, local initiatives, researchers, and authorities. The map provides everyone who is interested in the field access to this information, offering the possibility for science education and public engagement. The website is fed updated data and has been shared with stakeholders working in the field to make the field of renewable energy in Central and Eastern Europe more accessible.

2.3 Dialogical Formats

Many Social Labs of the NewHoRRIzon project came up with public engagement activities which in various ways provided spaces and methods to engage the public in dialogue research and innovation.

Quadralogue. The Quadralogue addresses barriers of communication and routine between individuals with different roles in research and innovation. By bringing together

[2] https://reknowhere.eu/ (Downloaded 11.02.2023) [28].

these individuals, who are otherwise not typically incentivized to discuss the bigger picture aspects of science and research with each other, the Quadralogue seeks to overcome this barrier to discuss the social impact of research and innovation. The design of the Quadralogue is a structured and facilitated 45-min dialogue-game.

By providing a unique 'gamified' environment to foster these conversations, the pilot action is a low-threshold way to bring together to share their expertise, concerns, experiences, and assumptions taken for granted in their normal day to day routine. The barriers are removed by the protocol of the game, as each of the four participants are responsible for sharing their interpretations and first impressions of the experiences from another participant's perspective. Another aspect is that the protocol encourages them to discuss with each other in plain, non-specialized language. Both an instructional video[3] explaining how to implement the Qaudralogue and a video[4] of an on-campus experience using the format at Ben Gurion University in Beersheba, Israel, have been shared to the public on YouTube [29, 30].

GenVoice. The ambition of this pilot action was to experiment with integrating the "unheard voices" of future generations who typically are not engaged, or involved in research only as future beneficiaries, into transport R&I processes – also in context of the contemporary civil society movements. The main target group was young adults who were invited to participate in an experimental workshop.

The morning session of GenVoice involved a school class of 16–17-year-old participants; the afternoon session, students 20–25 years old. The event followed a three-step process. First, participants debated about their personal experiences with transport in the area in Zilina (Slovakia), talked about their expectations for this workshop and described the travel experiences they make in their everyday lives. Second, they created visions of a desirable future and an ideal present mobility system. Third, solutions were created on how to make these visions become reality.

The pilot action left a lasting impression on both the participants and the organizers, in this case transport researchers. Specifically, the school class clearly enjoyed the openness of the process and being asked their opinion about contemporary issues, while also being able to bring in their own everyday experiences. They found the workshop fun, inspiring and empowering – and enjoyed being creative.

The organizers were satisfied as well, having left an impression on the participants both in terms of content and inducing a feeling of agency through eye-level conversations. Furthermore, the organizers pressed to put the results of the event on the radar of local policymakers and city-planners.

Research Goes to Streets was developed to address the lack of intellectual and physical connection between the academic universe and civil society. The format manifested into a one-day walkshop in Madrid, Spain, where participants attended in various stages. Although the specific topic addressed in the piloting of Research Goes to Streets was mobility and transportation, the format can be adapted to any topic which addresses

[3] https://www.youtube.com/watch?v=hXLWokWF7jU.
[4] https://www.youtube.com/watch?v=jqYcPmQvMRI.

a variety of stakeholder groups that otherwise are disconnected in their day to day personal or professional lives. For example, the initial pilot case included researchers, city technicians, students, mobility consultants, and members of diverse grassroots and NGOs.

Practically, the event began with two hours of promenade with several (four - five) stops. This stage included presentations from scientific experts on their research about the shared context, in this case the local transport system in that community, followed by direct questions from the attendants who were encouraged to bring in their own points of view and experiences. A video of the pilot implementation in Madrid is available on YouTube [31].

By bringing such a format out into the street, this pilot demonstrates the possibilities of dialogue between stakeholders and researchers on an eye-level, in a neutral space, outside, where community members can easily access. To extend the event further, as was done in the case of the pilot, there is also the option to record and make the event into a short film available to a wider audience.

2.4 Training Materials

Another group of pilots worked on various training materials which were informed by experiences and lessons from experimenting with public engagement approaches.

Knowledge Kiosk. This pilot was made in a series of co-creation workshops with the aim of using Design Thinking methodologies to develop an original and effective dialogue system between citizens and researchers to be sustained over a longer period of time. The inspiration behind the pilot was the observation that researchers who would like to contribute to public engagement are often not sure how to bring it into practice and lack examples of effective practices. The Knowledge Kiosk serves like a training manual and facilitation process for designing a long-term dialogue format that is suitable to local circumstances and context, fostering two-way dialogue along the way.

The first step of the Knowledge Kiosk exclusively targeted citizens who provided their ideas for how a regular, sustained interaction between citizens and scientists might look like. The second round was for scientists to discuss and develop these ideas further. Lastly, the groups are brought together to develop a prototype to fit both their needs and desires.

Good Practices of Co-creation. This pilot action is an example of a public engagement training developed in the specific context of healthcare. In this sense, the 'public' addressed by the pilot is patients and the specific issue is the disconnect between them, healthcare providers, industry, researchers, and policy makers. One result of this is a growing sense that patients' healthcare needs and wishes are not always properly met.

Co-creation is seen as an approach that can help to reduce this disconnect and strengthen the role of patients and relatives in health care research.

Various positive examples of co-creation exist in healthcare policy making, research, product and service development as well as clinical decision making. The pilot action

identified suitable co-creation initiatives and interviewed a few of them. The pilot identified the need to spread knowledge about these positive examples through "Co-Creation examples" that show the benefits of co-creation in health. The training approach taken in this pilot involves starting by broadening the horizons of affected stakeholders and improving awareness of co-creation initiatives and their benefits. The results of this pilot action can help those interested in co-creating feasible, acceptable, and effective healthcare processes.

Public Engagement from "Nice to Have" to "Need to Have". This pilot on public engagement was developed to increase the participation of citizens and stakeholders in social and technological innovations related to sustainable development. The growing public impatience around implementing sustainability contrasts with the simultaneous backlash from other parts of the public that do not feel represented by proponents of rapid societal transformations.

To help alleviate these tensions and obstacles to collective, inclusive, and sustainable innovation, the pilot action sought to collect and develop clear arguments for why public engagement is important in environmental research and innovation.

A survey was conducted amongst members of business, research, civil society, public officials and their networks. The questions used were designed to inform arguments that can be used to convince funding agencies and project partners about the necessity of public engagement and, in contrast, arguments against engagement. These arguments can be used in a variety of ways, one of which is to support organizations in their training efforts as the arguments represent the most common motivators and demotivators of public engagement activities. Some of these include (Table 1):

Table 1. Arguments for and against public engagement collected in "Public Engagement from 'Nice to Have' to 'Need to Have'".

Arguments for	Arguments against
• For producing findings/solutions/policies that are, on the long-term, acknowledged by a broad variety of actors • Diversity of thoughts leads to better Research and Innovation outcomes • The public is going to be engaged anyway, do you want to be there? Or miss out? • Cultural entitlement, informed citizenry, young people and education and their empowerment	• Not necessary, confusing, expensive, time-consuming, would not deliver the right result • It is lengthy, people are inactive, it takes lots of additional resources • I don't want to share my idea openly before I have finalized it • Difficulties with recruitment

The result is a summary of arguments and experiences that can be used for a training incentive to reflect on the importance of public engagement in R&I calls and proposals.

2.5 Hands on Training

In addition to training materials, some of the pilot actions also worked on developing hands-on training that can be used, adapted to and realized in different contexts, with different topics and for different public engagement needs.

Training on Stakeholder Integration. This training pilot specifically addresses consortium lead partners and participants of research and innovation projects and incentivizes reflection on the value of stakeholder integration. The training includes hands-on, participatory exercises to engage stakeholders with supplemental case-studies highlighting the necessary skills for effective engagement.

During the development of the pilot, the team in charge provided a training opportunity and best practice example of public engagement for grant applicants and project leads. The training focused on the benefits of stakeholder integration to research especially as far as quality of outcomes and societal impact is concerned. The aim was to further this into regular training for specific target groups on the national and European level.

The design and piloting of the training was led by an expert in multi-stakeholder processes and took place in February 2020 as a 1-day workshop in Vienna. A diverse group participated including researchers and representatives from national and European level funding organizations. A theoretical and practical insight into multi-stakeholder processes followed by applying a case-study in several steps of intensity of integration. Barriers and opportunities became obvious and were reflected on in the last part of the training.

2.6 Capacity Building

A final category of pilots relates to assisting individuals and organizations in their capacity to conceptualize, plan and implement appropriate public engagement exercises for their unique needs. These pilot actions demonstrate different approaches to transforming structures and systems to support public engagement practices and can be transferred and adapted to new organisations.

RRI Career Assessment Matrix. This pilot addresses research careers and how, in many research contexts, they are evaluated based on narrow definitions of excellence. The problem that this pilot addresses is that by evaluating successful careers through narrow lenses and leaving out public engagement practices, career assessment ultimately restricts diversity in academia, hindering its labor force and its approaches for addressing societal challenges.

Thus, the Social Lab envisioned a matrix as a means of change in the current evaluation frameworks and practices. The development of the Matrix involved a plenary session and participatory workshop during the Marie Curie Alumni Association Conference in February 2019 in Vienna. Based on this input, a policy brief was developed titled, “Towards Responsible Research Career Assessment” [32]. In the brief are five recommendations including a call to MSCA policymakers to broaden current evaluation

criteria of MSCA calls in dialogue with all relevant stakeholders. Other recommendations include current developments in both indicator development and narrative evaluation. Some examples of the core elements of such a matrix as presented in the policy brief [32] are as follows in Table 2.

More broadly, the pilot helps encourage funding institutions and research performing organizations to rethink and adapt institutional assessment and reward structures from a responsibility perspective. This means including elements like public engagement, teaching, and community service as an equally legitimate and rewarding cause for a researcher. Other organizations could use the policy brief, its sources and the process underlying it as an inspiration for improving their career evaluation system. The high-level policy brief was embraced by the Marie Curie Alumni Association.

Table 2. Excerpt of recommendations from "Towards Responsible Research Career Assessment" policy brief.

Recommendations	Dimensions
Broaden current evaluation criteria of Marie Skłodowska-Curie Action (MSCA) calls in dialogue with all relevant stakeholders	Robustness: basing metrics on the best possible data in terms of accuracy and scope; -
Provide (online) training for evaluators on implicit bias	Reflexivity: recognising and anticipating the systemic and potential effects of indicators, and updating them in response."
Offer training within the MSCA programme, such as via Innovative Training Networks, to prepare researchers and organizations for open and responsible, academic as well as non-academic careers	Humility: recognising that quantitative evaluation should support – but not supplant –qualitative, expert assessment;
Reward and showcase MSCA grantees who excel in multiple dimensions of research, teaching, and service	Diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system;
Support knowledge exchange and communities of practice around diverse and inclusive forms of excellence	Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results;

Measuring the Impact of RRI. This pilot addressed the topic of measuring the impacts of RRI at project level. An important driver for the pilot was to be able to easily share the findings with non-academic and academic audiences. The outcome was an easy-to-use template that can support a wide range of stakeholders in their evaluation of RRI activities, including public engagement. The first version of the template includes a list of economic, democratic, and societal indicator descriptions based on pre-existing MoRRI indicators, some examples of which are shared in Table 3 below [33].

The potential users of the template are researchers, practitioners and particularly stakeholders who are involved in research projects and would like to demonstrate how their participation has had an impact outside of the research community. Thus, the pilot action helps support the development and emergence of good practices by monitoring and demonstrating how RRI activities such as public engagement can enrich research and innovation contexts.

The indicators can be elaborated on collaboratively with researchers and stakeholders in their own context with discipline-specific experiences and expertise. This pilot also increases awareness about the need for these types of monitoring and evaluation, utilities, and deepening work for future practical contexts.

Table 3. Example indicators taken from "The Impact of RRI Template".

Indicators	Scientific impacts/benefits of RRI	Economic impacts/benefits of RRI	Societal and democratic impacts/benefits of RRI
Short-term impacts	Increased collaboration with other sectors (industry, public sector, civil society…)	Relationship building between previously siloed sectors	Evidence on the positive effects of science education
Medium-term impacts	Proactive outreach and engagement activities with previously siloed actors in society	Market rewards will favour institutions with leadership that promotes ethical and responsible relationship between science, society, and economy	Evidence on the positive effects of science education
Long-term impacts	Diversifying the pool of researchers (this will impact the diversity of knowledge)	Alignment of normative standpoint on impact goals and mitigation of negative impacts	Improved education system

The Future of Science? Society. This pilot action helped to address visions (or lack thereof) of a European research landscape that is societally engaged. At the core of these visions is the uncertainty about the future role of Responsible Research and Innovation (RRI) and the Science with and for Society (SwafS) programme, both which were promoted in Horizon 2020. Together, stakeholders and supporters of a new and advanced SwafS-like programme developed scenarios of multiple, plausible futures of science-society interactions.

The realization of these visions resulted in three different actions. First, the pilot action contributed to the "Pathways declaration to support RRI in the Horizon Europe" and established links to further SwafS projects as signatories for the declaration. Second, the pilot action engaged with others in the NewHoRRIzon project to mobilize SwafS stakeholders to take part in the public consultation process on Horizon Europe.

Finally, the pilot action performed a highly interactive scenario workshop with stakeholders who, guided by a thorough methodology, created the four different scenarios of the political, societal and research landscape in 2038 in the European Union. These novel scenarios represent the product of profound discussions and evaluations of a wide range of political, societal, economic, technological, and ideological factors and variables that might evolve very differently and alter the course of science-society relations. The four scenarios can be found in full in a journal article [34] and can be used to discuss challenges and opportunities related to the different political and ideological paradigms predominating four radically different future relationships between science and society.

3 Discussion

Social Lab participants, starting from their needs and ideas, came up with very different pilot actions in the context of their situation.

Some pilot actions addressed the knowledge gap between experts and laypeople and the need to inform the public about research and innovation and RPOs in a one-way communication and thus increase interest and chances of transferability of what they are doing.

Others came up with dialogical formats which addressed the general public or specific segments of the public, e.g., young people and students. The sites they used transformed streets, neighborhoods, organizations and most importantly, dialogical norms between researchers, non-researchers, community members and administrators.

The experimentation with the Social Labs within the NewHoRRIzon project also showed that creating formats for public engagement is not sufficient; it needs actual training, training materials, exchange, and knowledge-transfer to show researchers how to engage with the public.

Formats, training materials and training are still not enough to promote public engagement. The public engagement pilots developed in NewHoRRIzon also highlight that institutional structures are needed so that public engagement activities add to researchers' careers and are not a burdensome add on or even an obstacle to their career. To reduce these potential burdens, there is also the need to measure the output and outcome of public engagement activities. However, whether public engagement in research and innovation has a future also depends very strongly on how the relationship between science and society is perceived.

The pilot actions generated in the NewHoRRIzon project in a bottom up, experimental, and experiential approach showed that there is a strong call for public engagement activities from the stakeholders. They showed that such activities can address the general public and at the same time, very specific segments of society with very different formats and spaces. They showed that public engagement activities must be embedded in a nourishing research landscape that systematizes, exchanges, trains and provides the

institutional and procedural preconditions so that public engagement in research and innovation can flourish. The pilots also showed that a societal discourse about research, innovation and societal transformation is needed and that it is perceived that research and innovation are democratic endeavors in which the public has its active and rightful place.

Acknowledgements. The NewHoRRIzon research project on which this chapter is based was carried out from 2017 to 2021. It was funded by the European Commission within the Science with and for Society Programme (SwafS) with 6.8 Mio Euro (Grant agreement ID: 741402). We are grateful for funding and also thankful to all our NewHoRRIzon partners who put together, facilitated, kept going and motivated the Social Labs. We are particularly grateful to all the stakeholders who engaged in NewHoRRIzon as Social Lab participants and contributed their skills, perspectives, and engagement to the workshops. Finally, the authors would like to emphasize that these participants created the pilot actions presented in this paper. More information about the project and its results are available on www.newhorrizon.eu and https://cordis.europa.eu/project/id/741402.

References

1. Dewy, J.: Search for the public. In: Rogers, M.L. (ed.) The Public and Its Problems: An Essay in Political Inquiry, pp. 41–59. Penn State University Press (2012)
2. Bensaude-Vincent, B.: A historical perspective on science and its "others." Isis **100**, 359–368 (2009). https://doi.org/10.1086/599547
3. Jasanoff, S.: Technologies of humility: citizen participation in governing science. Minerva **41**, 223–244 (2003). https://doi.org/10.1038/450033a
4. Huttunen, S., Ojanen, M., Ott, A., Saarikosk, H.: What about citizens? A literature review of citizen engagement in sustainability transitions research. Energy Res. Soc. Sci. **91**, 102714 (2022). https://doi.org/10.1016/j.erss.2022.102714
5. Niels, M.: Science's disparate responsibilities: patterns across European countries. Public Underst. Sci. **27**(3) (2017). https://doi.org/10.1177/09636625177246
6. Bensaude Vincent, B.: The politics of buzzwords at the interface of technoscience, market and society: the case of "public engagement in science." Public Underst. Sci. **23**(3), 238–253 (2014). https://doi.org/10.1177/0963662513515371
7. Jahel, M., Hannah, V., Ellenbeck, S., Milan, B.F., Jaeger, C.: Stakeholder involvement in sustainability science—a critical view. Energy Res. Soc. Sci. **17**, 71–81 (2016). https://doi.org/10.1073/pnas.1706085114
8. Chilvers, J., Kearnes, M.: Remaking participation in science and democracy. Sci. Technol. Hum. Values **45**(3), 347–380 (2020). https://doi.org/10.1177/0162243919850885
9. Loet, L., Ward, J.: Science shops: a kaleidoscope of science-society collaborations in Europe. Public Underst. Sci. **14**, 353–372 (2005). https://doi.org/10.48550/arXiv.0911.4289
10. Horst, M., Michael, M.: On the shoulders of idiots: re-thinking science communication as "event." Sci. Cult. **20**(3), 283–306 (2011). https://doi.org/10.1080/09505431.2010.524199
11. Horst, M., Irwin, A.: Nations at ease with radical knowledge. On consensus, consensusing and false consensus. Soc. Stud. Sci. **40**(1), 105–126 (2010). https://doi.org/10.1177/0306312709341500
12. Irwin, A., Jensen, T.E., Jones, K.E.: The good, the bad and the perfect: criticizing engagement practices. Soc. Stud. Sci. **43**(1), 118–135 (2012). https://doi.org/10.1177/0306312712462461

13. Laurent, B.: Technologies of democracy: experiments and demonstrations. Sci. Eng. Ethics **17**, 649–666 (2008). https://doi.org/10.1007/s11948-011-9303-1
14. Rome Declaration on Responsible Research and Innovation in Europe. European Commission, Rome (2014). https://digital-strategy.ec.europa.eu/en/library/rome-declaration-responsible-research-and-innovation-europe. Accessed 7 July 2021, 19 Mar 2023
15. Hassan, Z.: The Social Labs Revolution. A New Approach to Solving Our Most Complex Challenges. San Francisco: Berret-Koehler Publishers, Inc. (2014). https://doi.org/10.48558/p3xv-km43
16. Timmermans, J., Blok, V., Braun, R., Wesselink, R., Nielsen, R.Ø.: Social labs as an inclusive methodology to implement and study social change: the case of responsible research and innovation. J. Responsib. Innov. **7**(3), 1–17 (2020). https://doi.org/10.1080/23299460.2020.1787751
17. Braun, R., et al.: Social labs as temporary intermediary learning organizations to help implement complex normative policies. Case Responsib. Res. Innov. Eur. Sci. Gov. (2022). https://doi.org/10.1108/TLO-09-2021-0118
18. Marschalek, I., Seebacher, L.M., Unterfrauner, E., Handler, K., Hofer, M.: Social labs in public engagement processes of RRI. In: Blok, V. (ed.) Putting Responsible Research and Innovation into Practice, Library of Ethics and Applied Philosophy, vol. 40, pp. 147–169 (2023). https://doi.org/10.1007/978-3-031-14710-4_8
19. Griessler, E., Blok, V.: Conclusion: implementation of responsible research and innovation by social labs. lessons from the micro-, meso- and macro perspective. In: Blok, V. (ed.) Putting Responsible Research and Innovation into Practice, Library of Ethics and Applied Philosophy, vol. 40 (2022). https://doi.org/10.1007/978-3-031-14710-4_14
20. Griessler, E., Hönigmayer, H.: D2.5 Actions and Activities to Realize RRI in Excellent Science (2021a). https://cordis.europa.eu/project/id/741402/results
21. Griessler, E., Hönigmayer, H.: D3.5 Actions and Activities to Realize RRI in Industrial Leadership (2021b). https://cordis.europa.eu/project/id/741402/results
22. Griessler, E., Hönigmayer, H.: D5.5 Actions and Activities to Realize RRI in Diversity of Approaches. (2021c). https://cordis.europa.eu/project/id/741402/results
23. Daimer, St., Yorulmaz, M., Griessler, E., Hönigmayer, H.: D4.5 Actions and Activities to Realize RRI in Societal Challenges (2021). https://cordis.europa.eu/project/id/741402/results
24. Cohen, J.B., Loeber, A.: Changing the research and innovation system through democratic experimentation. A guide to good practices for Responsible Research and Innovation, Amsterdam (2021). https://newhorrizon.eu/wp-content/uploads/2021/12/NH_D8.3_Guide_to_Good_Practices_FINAL_publicv2.pdf
25. NewHoRRIzon (w.d.) RRI Booklet. Unpublished brochure. https://newhorrizon.wpengine powered.com/wp-content/uploads/2022/04/NewHoRRIzon-Pilot-Action-Booklet_fin20220412.pdf
26. Jacobsen, B., Wallinger, M., Kobourov, S., Nöllenburg, M.: MetroSets: Visualizing sets as metro maps. IEEE Trans. Vis. Comput. Graph. (2020). https://arxiv.org/abs/2008.09367 to appear in InfoVis 2020
27. Euro-Expert & RRI. https://euro-resp.com
28. Renewable Energy Knowhere. https://reknowhere.eu/
29. Lewis, E.: Guided Video. https://www.youtube.com/watch?v=hXLWokWF7jU
30. Lewis, E.: Quadralogue BGU 2019. https://www.youtube.com/watch?v=jqYcPmQvMRI
31. https://www.youtube.com/watch?v=lFPsWMHyyHU
32. Cohen, J.B., et al.: Towards responsible research career assessment. Zenodo (2019). https://doi.org/10.5281/zenodo.3560479
33. NewHoRRIzon (w.d.) The Impact of RRI Template. Unpublished template

34. Daimer, St., Havas, A., Kuhls, K., Yorulmaz, M., Vrgovic, P.: Multiple futures for society, research, and innovation in the European Union: Jumping ahead to 2038. J. Responsib. Innov. **8**(2), 148–174 (2021). https://doi.org/10.1080/23299460.2021.1978692

Looking into the Future: Main Challenges

Evidence-Based Research Integrity

Ana Marušić(✉)

Center for Evidence-Based Medicine, University of Split School of Medicine, Šoltanska 2, 21000 Split, Croatia
ana.marusic@mefst.hr

Abstract. This chapter discusses the challenges to research integrity from the perspective of good research practices as defined by the European Code of Conduct for Research Integrity: 1) research environment, 2) training, supervision and mentoring, 3) research procedures, 4) safeguards, 5) data practices and management, 6) collaborative working, 7) publication and dissemination, and 8) reviewing, evaluating and editing.

Keywords: research integrity · research ethics · European Code of Conduct for Research Integrity · good research practices · research misconduct · detrimental research practices

1 Introduction

Research integrity is a rapidly expanding field of research, with the number of articles published in the last decade (2010 to 2022) almost three times greater than in each previous decade [1, 2]. This chapter will describe the latest evidence we have from research on research integrity. It is not a formal and methodologically rigorous scoping review, but a personal perspective from a meta-researcher in the European research landscape. The chapter is also skewed towards biomedical research, as that is the discipline I work in and study.

It is important to first provide the definition of research integrity, as it overlaps with other concepts – research ethics, responsible research and innovation (RRI), an open and responsible science.

From the perspective of the European research framework (ENERI – European Network for Research Ethics and Integrity), research integrity is defined as "the attitude and habit of the researchers to conduct research according to appropriate ethical, legal, and professional frameworks, obligations and standards" [3]. Research ethics is considered to be a wider concept, defined as "the application of ethical principles or values to the various issues and fields of research. This includes ethical aspects of the design and conduct of research, the way human participants or animals within research projects are treated, whether research results may be misused for criminal purposes, and it refers also to aspects of scientific misconduct" [4].

Research ethics (and integrity) is included in the wider concept that aligns research with the society – responsible research and innovation (RRI). This societal concept of

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 173–187, 2023.
https://doi.org/10.1007/978-3-031-33177-0_11

research is particularly emphasized in the research frameworks of the European Union, and is often defined as "a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process" [5]. In the 2020–2024 strategy on research and innovation of the European Commission [6], the goal related to the European Union (EU) Open Science Policy includes relevant concepts, such as research integrity and reproducibility of scientific results and citizen science, and the term open and responsible research/science is often used in the grant calls in Horizon Europe.

In addition to the differences in definitions and concepts, the European landscape is also very varied in the approaches to and formal structure of research ethics and research integrity, as demonstrated by our latest study [7], drawing from the Mutual Learning Exercise (MLE) on Research Integrity [8] and the information from the Embassy of Good Science [9]. We created Country Research Integrity Cards, describing the structures, processes and incentives for research integrity in several European countries; they are available at The Embassy of Good Science, a MediaWiki platform research ethics and integrity for the research community. Despite quite large differences in how different European countries legislate and structure research integrity, they are all unified around the core principles outlined in the European Code of Conduct for Research Integrity [10]: reliability, honesty, respect and accountability. The Code has become a kind of "soft" law in the EU, because all legal entities participating in the Horizon Europe research framework programme have to confirm that their activities will comply with the Code [11].

According to the Code, research integrity and responsible research is achieved through good research practices related to the 8 main contexts: 1) research environment, 2) training, supervision and mentoring, 3) research procedures, 4) safeguards, 5) data practices and management, 6) collaborative working, 7) publication and dissemination, and 8) reviewing, evaluating and editing. I am going to present most recent evidence, as well as our experience in improving these aspects of research, and point to future challenges in research integrity policies and practices.

2 Research Environment

The evidence has moved from the view of research misconduct as an individual failure (a "rotten apple") to the responsibility of the institutions and the importance of environment for responsible research ("basket of apples") [12]. Good research practices in research organizations recommended by the Code include promoting awareness about research integrity and building up research culture. This includes creating research policies and procedures, as well as establishing the process for addressing and investigating allegations of research misconduct. To achieve this, the organizations must have adequate infrastructure and reward open and reproducible research in job promotions or hiring new researchers, particularly early career researchers.

These recommendations are clear and reasonable. However, there is little high-quality evidence about interventions to achieve a research climate that promotes research integrity. We recently performed a coping review of interventions to change organizational climate or culture in academic or research settings [13]. The terms "culture" and

"climate" are often used interchangeably, although the differences have been described in literature [13]. Research climate is usually considered to be the shared perception of researchers about research policies, practices and procedures, and research behaviours perceived as rewarding. On the other hand, research culture is a wider and more complex system of most prevalent basic norms, deep principles and shared opinions in research environment. The methodological approaches to measuring research climate and culture also differ – climate is usually measured by questionnaire surveys, whereas culture can be best measured using qualitative study designs.

In our systematic review of more than 32 thousand articles retrieved from five bibliographical databases by specific search strategy and a manual search of a number of grey literature resources (Clinicaltrials.gov, Open Science Framework, Prospero database, Basesearch.net, Google Scholar, Opengrey.org, Campbell Collaboration Library and Science.gov databases), we identified only 7 studies that tested interventions for organizational (not necessarily research) climate or culture. Six of the seven studies reported positive changes after the intervention. These changes were measured by repeated questionnaire surveys or by narrative reports after the intervention. As the methodology in these studies was low, it is not possible at the moment to provide recommendations to institutions. There are many different practices that organizations for performing research and those that fund research have to promote research integrity [14], but also many different factors that influence promotion and implementation of research integrity policies at these organizations [15]. More research and rigorous methodology are needed if we want to understand how interventions work at the institutional level and whether they are effective. A recent set of tools for generating research integrity promotion plans at research performing and research funding organizations, created by the Horizon 2020 project SOPs4RI (Standard Operating Procedures for Research Integrity) [16], may be a good start for testing research integrity interventions in organizations.

It is important to keep in mind that different organizations need different approaches as well as focus on different research integrity issues, especially in relation to the research disciplines involved, as they may have different research integrity standards and practices [17]. Furthermore, it seems that it is difficult to change attitudes, knowledge and behaviours in research integrity. A ten-year follow-up study of doctoral students at the Faculty of Medicine at the University of Oslo in Norway showed that, although there is an improvement in their research integrity, the research integrity indicators remained stable over time [18]. It has to be kept in mind that Norway has a well-developed research integrity and research ethics framework at national and organizational levels [7].

3 Training, Supervision and Mentoring

In order to achieve high standards of responsible research, education and training of researchers of all professional levels is mandatory, from early career researchers to senior researchers and managers. Educational activities should include not only ethics and integrity but also research design, methodology and analysis, so that they produce high quality and reproducible research.

Some years ago, we conducted a systematic review using rigorous Cochrane methodology to assess the evidence for interventions to prevent research misconduct and foster

research integrity [19]. We looked at 31 studies, including randomized controlled trials as the most rigorous methodological study design and involving more than 9 thousand participants. Like with organizational culture/climate, we do not have good evidence of what works, except for some low-quality evidence that indicates that practical teaching about plagiarism may reduce its occurrence in the student population.

It is thus very difficult to provide recommendations to research organizations about how to organize research integrity training. A recent content analysis of available training materials at 11 research-intensive European universities showed great diversity in the materials among the universities [20]. Recommendations that emerged from this study were that universities have to 1) develop university-wide research integrity training, at least at the level of postgraduate (doctoral) studies; 2) agree on the minimum requirements for the content, format, length and frequency of research integrity education; and 3) share educational material to increase the quality of research in whole Europe and facilitate mobility of researchers. Findings from a large focus study involving 147 participants in 8 European countries [21] support these recommendations. The themes that emerged from this qualitative study included the need for research integrity education to be available to all actors in research, to be tailored to specific needs and to use formal and informal format, and to be paralleled by active motivation of the trainees.

While there is a plethora of educational material for research integrity available online, especially in the USA [22], there was no systematic approach in Europe. In recent years, different approaches to teaching research integrity have been developed by several Horizon 2020 projects, including VIRT2UE, Path2Integrity, and Integrity. They came together in a new initiative, the Network for education in research quality (NERQ), which has the ambition to share good practices in teaching research integrity and generally high-quality research, improve the training of trainers and stimulate the development of new trainings based on evidence [23].

Of particular importance in promoting research integrity is supervision and mentoring, which can impart important values and virtues during professional development of a researcher [24]. In academic medicine, we showed some time ago that, while mentoring is perceived as very important, there is not much evidence on its effectiveness as an intervention tool to help the professional development of the mentees [25]. The scoping review of qualitative studies in academic mentoring in medicine also showed that both the mentor and the mentee had to be engaged in and committed to successful mentoring and the mentoring has to happen in a facilitating organizational environment [26]. The important role of the organization in supporting mentoring and supervision of early career researchers, particularly in relation to research integrity was confirmed in a recent qualitative study of research supervisors and their role of research integrity trainers [27]. More studies are needed to explore what interventions work to promote and facilitate successful mentoring. As senior researchers often do not succeed in transmitting the knowledge, behaviours and virtues necessary for responsible research, there are calls for reverse mentoring, where the early career researchers act as mentors to their senior colleagues. In this way, which has been tested in practice in some fields, it may be possible to more effectively promote and foster research integrity and build a positive climate for responsible research [28].

Mentoring, as a very personal and complex professional relationship between two individuals currently faces not only the problems of evidence for its effectiveness but also important ethical dilemmas in the world of growing use of artificial intelligence (AI) in research, including mentoring [29]. It will be interesting to follow future developments in AI-assisted mentoring.

In order to achieve high standards of responsible research, education and training of researchers of all professional levels is mandatory, from early career researchers to senior researchers and managers. Educational activities should include not only ethics and integrity but also research design, methodology and analysis, so that they produce high quality and reproducible research.

4 Research Procedures

When performing research, researchers are expected to base their research on best available evidence; to use the resources to design and execute research, and analyse the results in a responsible way; interpret and publish the results in open and transparent, as well as accurate way so that they can be verified and reproduced.

Responsible approach to research procedures is an ongoing topic of debate and efforts to improve this important, pervasive and complex aspect of research. Many solutions have been offered in recent years but we still need more studies and better evidence to see what works to "increase the value and reduce waste in research design, conduct and analysis" [30].

In recent years, the term "reproducible science" has gained its importance as a point of intervention to increase the quality of research results. It includes different measures to assure the integrity and trustworthiness of important elements of scientific research: use of research methods, reporting and dissemination of results and their reproducibility, as well as evaluation of research and incentives for researchers [31]. Research reproducibility is one of the eight ambitions the EU's open science policy [6]. Table 1 presents the proposals to increase science reproducibility, with examples of interventions, however without evidence of their effectiveness as yet [31].

As can be seen from Table 1, most of the proposed initiatives had low uptake in 2017, and much has not been changed until the present. Some of the initiatives, such as PubMed Commons feature to promote post-publication peer review have failed, due to a lack of interest [32]. New modalities of reporting and dissemination have been developed, which I will address in the section on publishing.

Even those incentives that are considered to be widely adopted, such as registration of clinical trials, are not proving to be effective. For example, a study of registration status of published clinical trials showed that the prevalence of published trials that were pre-registered increased over time, but still remain very low, with only one in five trials being prospectively registered [33]. We still have a long time and more effort needed to address all aspects of quality research and test interventions to increase it.

Table 1. A manifesto for reproducible science.*

Theme	Proposal	Examples of initiatives/potential solutions	Extent to current adoption (in 2017)	Stakeholders
Methods	Protecting against cognitive biases	All of the initiatives listed below		
		Blinding	5–30%	Journals, Funders
	Improving methodological training	Rigorous training in statistics and research methods for future researchers	<5%	Institutions, Funders
		Rigorous continuing education in statistics and methods for researchers	<5%	Institutions, Funders
	Independent methodological support	Involvement of methodologists in research	5–30%	Funders
		Independent oversight	<5%	Funders
	Collaboration and team science	Multi-site studies/distributed data collection	<5%	Institutions, Funders
		Team-science consortia	<5%	Institutions, Funders
Reporting and dissemination	Promoting study pre-registration	Registered Reports	<5%	Journals, Funders
		Open Science Framework	<5%	Journals, Funders
	Improving the quality of reporting	Use of reporting checklists	5–30%	Journals
		Protocol checklists	<5%	Journals
	Protecting against conflicts of interest	Disclosure of conflicts of interest	30–60%	Journals
		Exclusion/containment of financial and non-financial conflicts of interest	<5%	Journals
Reproducibility	Encouraging transparency and open science	Open data, materials, software and so on	<5% to 5–30%	Journals, Funders, Regulators
		Pre-registration	>60% for clinical trials, <5% for other studies)	Journals, Funders, Regulators

(*continued*)

Table 1. (*continued*)

Theme	Proposal	Examples of initiatives/potential solutions	Extent to current adoption (in 2017)	Stakeholders
Evaluation	Diversifying peer review	Preprints	<5% in biomedical/behavioural sciences, >60% in physical sciences	Journals
		Pre- and post-publication peer review, for example, Publons, PubMed Commons	<5%	Journals
Incentives	Rewarding open and reproducible practices	Badges	<5%	Journals, Institutions, Funders
		Registered Reports	<5%	Journals, Institutions, Funders
		Transparency and Openness Promotion guidelines	<5%	Journals, Institutions, Funders
		Funding replication studies	<5%	Journals, Institutions, Funders
		Open science practices in hiring and promotion	<5%	Journals, Institutions, Funders

*Reproduced from [31] under Creative Commons Attribution 4.0 International Licence. The content of the table was not change, but the presentation was changed to include full description of the table legends in the table and adding additional column to visualise the adoption of initiatives in research practice.

5 Safeguards

This good research practice relates to compliance with codes and other professional regulations relevant to the research discipline, including the participation of people in research, or involvement of animals or cultural, biological or environmental research subjects. Research must be done with respect to the benefit of the community and not only of the researchers. Researchers must also plan, execute, analyse and report the results of their research with regard to differences among different populations, including age, gender, culture, personal beliefs, socioeconomic factors and ethnic origin. They must also address and manage not only benefits but also harms stemming from their research.

Based on historical principles and experiences, safeguards in science are often legally determined and well regulated, in order to anticipate and mitigate unintended consequences of scientific research, as defined by Merton already in 1936: ignorance, errors,

focusing on immediate benefit instead of long-term consequences, basic values, and self-defeating prophecy [34]. Despite long-existing codes and regulations, the prevalence of research misconduct, either in the form of fraud – falsification, fabrication and plagiarism, or detrimental research practices, which are smaller but very prevalent poor research practices, does not significantly changed over the last decade. A recent meta-analysis of reported practices estimated that about 3% or researchers report committing at least 1 fraudulent practice and about 13% at least one detrimental research practice [35]. About 16% of them have witnessed others committing fraud and 40% witnessed instances of detrimental research practices by others [35].

In the European context, there is a great diversity in available codes and other guidance documents, depending on the country, research discipline and the research development. A recent analysis of national level codes for research integrity [36], showed great divergence among European countries, despite the unifying umbrella of the European Code of Conduct for Research Integrity. What is most common are the codes to deal with scientific fraud, i.e. egregious research misconduct. Our analysis of the European landscape for research integrity in 2022 also demonstrated the lack of harmony in national structures, procedures and practices [7]. There is also diversity of how researchers from different European settings perceive research misconduct. A study of researchers from institutions in northern, southern and northwestern Europe had different perceptions of plagiarism practices, with those from more northern countries having a stricter view of what represents plagiarism [37].

All this presents a serious problem with regard to safety of researchers when they move between research organisations in different countries, because of uncertainty in expectations from research integrity codes. There is a lot of work at all levels, from the regulators to the research organizations and researchers to arrive to a common understanding of underlying principles of research integrity. The tools box for creating research integrity promotion plans from the Horizon 2020 project SOPs4RI, as mentioned before, may help the harmonization of codes across Europe [16, 38].

6 Data Practices and Management

Openness of data has become a central principle in science, particularly from the view of public research funders, who want to ensure a maximum use of research output. The FAIR principles of data management and stewardship are implemented in the European Open Science Policy [6]: research data should be findable, accessible, interoperable and reusable.

It is also expected from the actors in research (researchers and their organizations) to decide on how they will provide or allow the use of data and research materials. Data should also be acknowledged as product of research and their intellectual property should be protected.

There is not enough space in this chapter to address all aspects of research data practices. We have moved from physical laboratory books to online data laboratory diaries and large datasets as a standard research practice. This has brought great benefits, especially in research with “big data”, but has also created problems. Whereas the mismanagement of research data is considered a detrimental research practice, it is often

found in cases of research misconduct, which brought about the calls to consider data mismanagement as an act research misconduct, i.e. fraud, in some cases [39]. There are also calls to implement the FAIR principles not at the end of the research process, when the data gathering and analysis are complete, but throughout the research process, and for each data produced in individual experiments or studies [40].

We will have to follow the rapid developments in this field, especially the use of open repositories and specifically the European Open Science Cloud (EOSC) as a "trusted, virtual, federated environment … to store, share, process and reuse research digital objects (like publications, data and software" according to FAIR principles [6].

7 Collaborative Working

The European Code of Conduct for Research Integrity [10] requires researchers to take the responsibility for collaborative research, whether the research is between individual researchers, their institutions, across geographical or sector boundaries. Research collaboration is not only about the results of such research endeavour, but also about the accountability and openness at every stage of research, from its beginning. The collaborators should formally agree on the expectations and standards in collaborative research, on how the intellectual property of research outputs will be protected, and how conflicts or research misconduct will be addressed.

There is not much evidence on what works best for successful and responsible collaborative research, but there is guidance from international consultations on this topic. The Montreal Statement of Research Integrity in Cross-Boundary Research Collaborations [41] defines general collaborative responsibilities of researchers and their institutions (integrity; trust; purpose; and goals), their responsibilities in managing the collaboration (communication; agreements; compliance with laws, policies and regulations; costs and rewards; transparency; resource management; and monitoring), responsibilities in collaborative relationships (roles and responsibilities; customary practices and assumptions; conflict; and authority of representation), and responsibilities for the outcome of research (data, intellectual property and research records; publication; authorship and acknowledgment; responding to irresponsible research practices; and accountability). Although collaborative research is often viewed as a characteristic of biomedical and natural sciences, integrity is an important aspect of intercultural research in social sciences and humanities [42].

Research collaboration occur between sectors, which can raise specific problems. For example, it has been shown that researchers from the industry sector differ in their perception of research integrity from their colleagues at universities [43]. In research involving patient and public involvement in research, accountability is an important part of the framework for such collaborations [44] and should be actively implemented in citizen science, as one of the ambitions in the European open science policy [6].

Collaborations across different types of borders create special problems for investigating allegations of research misconduct because of rather large differences in research integrity bodies, regulations and procedures [7]. The OECD Global Science Forum has suggested a practical guide for "Investigating Research Misconduct Allegations in International Collaborative Research Projects" [45]. It calls for the promotion of generally

acceptable responsible research practices and compliance with the national law where the individual researcher is employed or is based, or where research takes place or where the facilities for the project may be located. Standard investigation procedures should include agreement on with who and where the responsibility for the investigation lies, and the procedure to be followed, with all parties providing assistance with the investigation.

8 Publication and Dissemination; Reviewing, Evaluating and Editing

I am addressing the last two good research practices from the European Code of Conduct for Research Integrity in a single section because they represent the two sides of the same concept: researchers publishing and communicating their work and researchers taking part in the evaluation of research results, be they reviewers of grant proposal, journal articles or research advancement request or journal editors.

As authors, researchers take the responsibility for their work presented in a publication, they agree on authorship of their works, and make their work available to the colleagues, including both positive and negative results. They also have to be honest when they communicate the results of their research to the media and the public, as well as they declare their activities and relationships that may create conflict of interest with regard to the submitted research. They should also acknowledge the contributions of individuals who have influenced their work, including accurate citations of previously published research. They are also responsible for correcting the published record of their work when it is necessary.

When they act as reviewers or evaluators of journal manuscripts, grant proposals or appointment, promotion or reward submission, they must do it in a responsible manner and without conflicts of interest. They have to maintain the confidentiality of the process when required and respect the rights of authors of the work they evaluate.

There is ample body of research, including interventions to improve different aspects of the publication and evaluation process, so I will here focus on some of the emerging issues, such as peer review, preprints and the use of AI in manuscript writing, and the "oldest" issue in publishing – authorship. Generally, we know from more than 30 years of research into peer review and scientific publication that this type of research is burdened by all challenges as other research field, including a significant publication bias, and lack of studies with rigorous methodology to test potential interventions [46].

With regard to authorship, my research group has performed a number of studies, including randomized controlled trials to demonstrate that different format of declaring authorship contributions do not work well for determining deserved authorship [19], and that the decisions on authorship are based on moral- rather than rule-based reasoning [47]. Based on our research, we believe that it is fairer to ask authors why they think they deserve to be authors on an article and publish this information [48], and that discussion about authorship should not occur at the end of the research study, but should be a part of the research protocol and monitored during the study [49]. A recent systematic review of the ethics of the Contributor Role Taxonomy (CRediT) system of declaring authorship in published articles also showed that the categorization of authorship contributions does not prevent unethical attribution of authorship [50] and that a structural transformation of

the scientific process and scholarly publication is needed to promote honest, responsible, and above all, deserving authorship.

Peer review has a central role in assuring the quality of published research, but we still do not know what is the best way to do it, either for journal articles [46] or grant proposals [51]. However, there is a lot of experimentation and innovation going on in the field of peer review, from traditional double- or single-blind peer review to open peer review, as well as results-free, consultative and post-publication peer review [52]. It will be interesting to see the results from the evaluation of these news practices.

Another innovation in scientific publishing are preprints – complete versions of a research manuscript that have not been peer reviewed [53]. Preprints are very common in some research discipline, especially physics, but the COVID-19 pandemic has introduced them into other fields, like medicine, which had been generally skeptical about the usefulness of preprints and aware of public health hazards from research that has not passed a quality control. However, it seems that preprints are here to stay, and we will need to answer many questions and resolve dilemmas about their use. The editorial and publishing community is working on ensuring that there is clear indexing and linking of preprints with the published versions of the article and that the responsibilities for the integrity of the published records are defined for all stakeholders [53].

It is fitting to close this section with the newest challenge in scientific publishing – the use of AI, i.e. computer programmes that can process human conversation and simulate it in communication with human. There are already several published articles where the AI (ChatGPT programme) was listed as a co-author [54]. Apart from philosophical question of sentience of a computer programme, the use of AI in scientific writing also raises research integrity issues, particularly with regard to deserving authorship, conflict of interest and potential plagiarism. Just as with preprints, it seems that AI is going to be increasingly used in writing research papers. However, at the moment it is a (very impressive and powerful) tool, but not an author. The current position of journal editors in biomedicine is that [54]: a) a ChatGPT cannot be an author because it cannot satisfy the current authorship criteria in biomedicine (approval of publication and taking accountability for content); b) it does not understand its conflict of interest and cannot legally sign a statement or hold copyright; c) (human) authors have to be transparent how and where AI tool was used in the paper; and d) (human) authors are responsible for the work performed by AI on their paper, including the accuracy of what is presented, including references cited, and the absence of plagiarism. The future for the publishing world will surely bring new tools for editors to detect parts of the papers written by AI [54].

9 Instead of a Conclusion

This chapter does not have a conclusion, because there are so many existing and rapidly emerging challenges to responsible research and integrity of researchers. We have to be aware that the future will bring some solutions, as well as new challenges. What we have to keep in mind is that, if we want research to be responsible and performed with integrity of all participants, we should look for facts and not follow authorities, and learn as much as we can so that we can make evidence-based policies that work for the benefit of the global community of humanity.

References

1. Marušić, S.L., Marušić, A.: Codes of ethics and research integrity. In: Laas, K., Davis, M., Hildt, E. (eds.) Codes of Ethics and Ethical Guidelines. TILELT, vol. 23, pp. 83–96. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-86201-5_5
2. Soehartono, A.M., Yu, L.G., Khor, K.A.: Essential signals in publication trends and collaboration patterns in global research integrity and research ethics (RIRE). Scientometrics **127**(12), 7487–7497 (2022). https://doi.org/10.1007/s11192-022-04400-y
3. ENERI – European Network for Research Ethics and Integrity. What is research Integrity? (2019). https://eneri.eu/what-is-research-integrity/
4. ENERI – European Network for Research Ethics and Integrity. What is research ethics? (2019b). http://eneri.eu/what-is-research-ethics/
5. von Schomberg, R.: Prospects for technology assessment in a framework of responsible research and innovation. In: Dusseldorp, M., Beecroft, R. (eds.) Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden. Vs Verlag, Wiesbaden (2011)
6. European Commission. Research and innovation strategy 2020–2024 (2020). https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024_en
7. Perković Paloš, A., Roje, R., Tomić, V., Marušić, A.: Creating research ethics and integrity country report cards: case study from Europe. Account. Res., 1–35 (2023). https://doi.org/10.1080/08989621.2022.2163632
8. European Commission.Mutual Learning Exercise (MLE) on Research Integrity (2018). https://ec.europa.eu/research-and-innovation/en/statistics/policy-support-facility/mle-research-integrity
9. Embassy of Good Science. Facts and Figures (2020). https://embassy.science/wiki/Special:BrowseData/Report
10. ALLEA – All European Academies: The European Code of Conduct for Research Integrity (2017). https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf
11. European Commission. Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013. The Official Journal of the European Union. L170/1 (2021). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R0695&from=EN
12. Bouter, L.M.: Commentary: perverse incentives or rotten apples? Account. Res. **22**(3), 148–161 (2015). https://doi.org/10.1080/08989621.2014.950253
13. Viđak, M., Barać, L., Tokalić, R., Buljan, I., Marušić, A.: Interventions for organizational climate and culture in academia: a scoping review. Sci. Eng. Ethics **27**(2), 1–23 (2021). https://doi.org/10.1007/s11948-021-00298-6
14. Ščepanović, R., Labib, K., Buljan, I., Tijdink, J., Marušić, A.: Practices for research integrity promotion in research performing organisations and research funding organisations: a scoping review. Sci. Eng. Ethics **27**(1), 1–20 (2021). https://doi.org/10.1007/s11948-021-00281-1
15. Roje, R., Reyes Elizondo, A., Kaltenbrunner, W., Buljan, I., Marušić, A.: Factors influencing the promotion and implementation of research integrity in research performing and research funding organizations: a scoping review. Account. Res., 1–39 (2022). https://doi.org/10.1080/08989621.2022.2073819
16. Horbach, S.P.J.M., Bouter, L.M., Gaskell, G., Hiney, M., Kavouras, P., Mejlgaard, N., et al.: Designing and implementing a research integrity promotion plan: recommendations for research funders. PLoS Biol. **20**(8), e3001773 (2022). https://doi.org/10.1371/journal.pbio.3001773

17. Sørensen, M.P., et al.: Strengthening research integrity: which topic areas should organisations focus on? Hum. Soc. Sci. Commun. **8**, 198 (2021). https://doi.org/10.1057/s41599-021-00874-y
18. Hofmann, B., Thoresen, M., Holm, S.: Research integrity attitudes and behaviors are difficult to alter: results from a ten-year follow-up study in Norway. J. Empir. Res. Hum. Res. Ethics, 15562646221150032 (2023). https://doi.org/10.1177/15562646221150032
19. Marušić, A., Wager, E., Utrobičić, A., Rothstein, H.R., Sambunjak, D.: Interventions to prevent misconduct and promote integrity in research and publication. Cochrane Database Sys. Rev. **4**(4), MR000038 (2016). https://doi.org/10.1002/14651858.MR000038.pub2
20. Abdi, S., Pizzolato, D., Nemery, B., Dierickx, K.: Educating PhD students in research integrity in Europe. Sci. Eng. Ethics **27**(1), 1–19 (2021). https://doi.org/10.1007/s11948-021-00290-0
21. Labib, K., et al.: Education and training policies for research integrity: insights from a focus group study. Sci. Pub. Policy **49**, 246–266 (2022). https://doi.org/10.1093/scipol/scab077
22. Pizzolato, D., Abdi, S., Dierickx, K.: Collecting and characterizing existing and freely accessible research integrity educational resources. Account. Res. **27**(4), 195–211 (2020). https://doi.org/10.1080/08989621.2020.1736571
23. NERQ – Network for education in research quality (2023). https://community.embassy.science/t/about-the-nerq-category/364
24. Tomić, V., Buljan, I., Marušić, A., VIRT2UE project consortium: development of consensus on essential virtues for ethics and research integrity training using a modified Delphi approach. Account. Res., 1–24 (2022). https://doi.org/10.1080/08989621.2022.2128340
25. Sambunjak, D., Straus, S.E., Marušić, A.: Mentoring in academic medicine: a systematic review. JAMA **296**(9), 1103–1115 (2006). https://doi.org/10.1001/jama.296.9.1103
26. Sambunjak, D., Straus, S.E., Marušić, A.: A systematic review of qualitative research on the meaning and characteristics of mentoring in academic medicine. J. Gen. Int. Med. **25**(1), 72–78 (2010). https://doi.org/10.1007/s11606-009-1165-8
27. Pizzolato, D., Dierickx, K.: Reverse mentoring to enhance research integrity climate. BMC Res. Notes **15**(1), 209 (2022). https://doi.org/10.1186/s13104-022-06098-w
28. Pizzolato, D., et al.: How can research institutions support responsible supervision and leadership? Account. Res., 1–23 (2022). https://doi.org/10.1080/08989621.2022.2112033
29. Köbis, L., Mehner, C.: Ethical questions raised by AI-supported mentoring in higher education. Front. Artif. Intell. **4**, 624050 (2021). https://doi.org/.3389/frai.2021.624050
30. Ioannidis, J.P., et al.: Increasing value and reducing waste in research design, conduct, and analysis. Lancet **383**(9912), 166–175 (2014). https://doi.org/10.1016/S0140-6736(13)62227-8
31. Munafò, M.R., et al.: A manifesto for reproducible science. Nat. Hum. Behav. **1**, 0021 (2017). https://doi.org/10.1038/s41562-016-0021
32. Retraction Watch. PubMed shuts down its comments feature, PubMed Commons (2018). https://retractionwatch.com/2018/02/02/pubmed-shuts-comments-feature-pubmed-commons/
33. Trinquart, L., Dunn, A.G., Bourgeois, F.T.: Registration of published randomized trials: a systematic review and meta-analysis. BMC Med. **16**(1), 173 (2018). https://doi.org/10.1186/s12916-018-1168-6
34. Merton, R.K.: The unanticipated consequences of purposive social action. Am. Soc. Rev. **1**(6), 894–904 (1936)
35. Xie, Y., Wang, K., Kong, Y.: Prevalence of research misconduct and questionable research practices: a systematic review and meta-analysis. Sci. Eng. Ethics **27**(4), 1–28 (2021). https://doi.org/10.1007/s11948-021-00314-9
36. Desmond, H., Dierickx, K.: Research integrity codes of conduct in Europe: understanding the divergences. Bioethics **35**(5), 414–428 (2021). https://doi.org/10.1111/bioe.12851

37. Yi, N., Nemery, B., Dierickx, K.: Do biomedical researchers differ in their perceptions of plagiarism across Europe? Findings from an online survey among leading universities. BMC Med. Ethics **23**(1), 78 (2022). https://doi.org/10.1186/s12910-022-00818-4
38. Mejlgaard, N., Bouter, L.M., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A., et al.: Research integrity: nine ways to move from talk to walk. Nature **586**(7829), 358–360 (2020). https://doi.org/10.1038/d41586-020-02847-8
39. Yeo-Teh, N.S.L., Tang, B.L.: Research data mismanagement - from questionable research practice to research misconduct. Account. Res., 1–8 (2023). https://doi.org/10.1080/08989621.2022.2157268
40. Dempsey, W., Foster, I., Fraser, S., Kesselman, C.: Sharing begins at home: how continuous and ubiquitous FAIRness can enhance research productivity and data reuse. Harv. Data Sci. Rev. **4**(3) (2022). https://doi.org/10.1162/99608f92.44d21b86
41. World Conference on Research Integrity. The Montreal Statement n Research Integrity in Cross-Boundary Research Collaborations (2013). https://www.wcrif.org/montreal-statement/file
42. Liu, X., Burnett, D.: Insider-outsider: methodological reflections on collaborative intercultural research. Humanit. Soc. Sci. Commun. **9**, 314 (2022). https://doi.org/10.1057/s41599-022-01336-9
43. Godecharle, S., Nemery, B., Dierickx, K.: Differing perceptions concerning research integrity between universities and industry: a qualitative study. Sci. Eng. Ethics **24**(5), 1421–1436 (2017). https://doi.org/10.1007/s11948-017-9965-4
44. Greenhalgh, T., et al.: Frameworks for supporting patient and public involvement in research: Systematic review and co-design pilot. Health Expect. **22**(4), 785–801 (2019). https://doi.org/10.1111/hex.12888
45. OECD Global Science Forum: Investigating Research Misconduct Allegations in International Collaborative Research Projects (2009). https://www.oecd.org/sti/inno/42770261.pdf
46. Malički, M., von Elm, E., Marušić, A.: Study design, publication outcome, and funding of research presented at international congresses on peer review and biomedical publication. JAMA **311**(10), 1065–1067 (2014). https://doi.org/10.1001/jama.2014.143
47. Hren, D., Sambunjak, D., Marušić, M., Marušić, A.: Medical students' decisions about authorship in disputable situations: intervention study. Sci. Eng. Ethics **19**(2), 641–651 (2013). https://doi.org/10.1007/s11948-012-9358-7
48. Malički, M., Jerončić, A., Marušić, M., Marušić, A.: Why do you think you should be the author on this manuscript? Analysis of open-ended responses of authors in a general medical journal. BMC Med. Res. Method. **12**, 189 (2012). https://doi.org/10.1186/1471-2288-12-189
49. Marušić, A., et al.: Five-step authorship framework to improve transparency in disclosing contributors to industry-sponsored clinical trial publications. BMC Med. **12**, 197 (2014). https://doi.org/10.1186/s12916-014-0197-z
50. Hosseini, M., Gordijn, B., Wafford, Q.E., Holmes, K.L.: A systematic scoping review of the ethics of contributor role ontologies and taxonomies. Account. Res., 1–28 (2023). https://doi.org/10.1080/08989621.2022.2161049
51. Pina, D.G., Buljan, I., Hren, D., Marušić, A.: A retrospective analysis of the peer review of more than 75,000 Marie Curie proposals between 2007 and 2018. eLife **10**, e59338 (2021). https://doi.org/10.7554/eLife.59338
52. Tennant, J.P., et al.: A multi-disciplinary perspective on emergent and future innovations in peer review [version 3; peer review: 2 approved]. F1000Research **6**, 1151 (2017). https://doi.org/10.12688/f1000research.12037.3
53. Stojanovski, J., Marušić, A.: Preprints are here to stay – is that good for science? In: Eaton, S.E. (ed.) Handbook of Academic Integrity, 2nd edn. Springer, Berlin (2023, in press)

54. Zielinski, C., et al.: For the WAME Board: Chatbots, ChatGPT, and Scholarly Manuscripts. WAME Recommendations on ChatGPT and Chatbots in Relation to Scholarly Publications (2023). https://wame.org/page3.php?id=106

Gender Mainstreaming and RRI: The Double Challenge

Santiago Garcia-Campa and Rosana Sanahuja(✉)

University Jaume I of Castellón, Castellón de la Plana, Spain
{gcampa,sanahujr}@uji.es

Abstract. Gender equality in research and innovation (R&I) has improved over the past two decades, although not without its ups and downs. The literature has pointed out the obstacles and difficulties in making gender equality effective in R&I while providing a wide range of strategies, methods and tools to achieve it. Responsible Research and Innovation (RRI) is a new approach that has made gender equality one of its six keys, raising a new challenge for gender mainstreaming. The Horizon Europe Framework Program 2021–2027 represents a step forward, with the Gender Equality Plan as a reference tool for mainstreaming gender in R&I organizations. However, there are more comprehensive proposals, such as the European Institute for Gender Equality's Gender Mainstreaming Platform. This chapter presents the platform's content and shows its usefulness in helping prepare Gender Equality Plans and promoting the institutional transformation of R&I organizations.

Keywords: Gender equality · Responsible Research and Innovation · Mainstreaming · Gender Equality Plan · European Institute for Gender Equality · Platform on Gender Mainstreaming · Open Science

1 Introduction

Gender equality has been becoming increasingly important ever since the mid-20th century. Following its recognition as a human right in the principal international and regional declarations of rights [1], academic and institutional interest has progressively focused on how to make it effective [2]. As a result of the contributions made for this purpose, we have both a wide range of strategies for making gender equality effective (gender mainstreaming, affirmative actions, gender balance, work-life balance, etc.) and a rich set of methodologies and tools designed to put them into practice (gender analysis, gender statistics and indicators, gender audits, gender budgeting, gender-responsive public procurement, and so on).

Research and Innovation (R&I) has also been concerned with gender equality, meeting obstacles and resistance similar to those encountered in other areas of public involvement [3]. This situation followed the pattern observed in many other areas, which have moved from gender blindness to making it relevant, mainly as a consequence of the

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 188–202, 2023.
https://doi.org/10.1007/978-3-031-33177-0_12

contributions of academic feminism [4]. However, this process has also had its ups and downs [5], interruptions [6] and even setbacks in periods of crisis [7].

Recently, the Responsible Research and Innovation model (hereinafter, RRI) promoted by the European Union, has conceived gender equality as one of its six key areas [8]. This approach gives a new direction to the relationship between gender equality and R&I, both because it departs from gender mainstreaming in the strict sense – that is, as a dimension that should be present in the other key areas – and because of its new implications for understanding R&I.

The purpose of this chapter is to contrast the reduction of gender equality to an area of RRI with its definition as a transversal dimension that must be integrated into any facet of R&I; find out the state of the relationship between gender equality and R&I, paying special attention to the most important challenges to be faced; showcase Gender Equality Plans as a leading tool chosen by the European Commission in its Horizon Program for funding R&I; and explain the main components and factors of the European Institute for Gender Equality's Gender Mainstreaming Platform, as they are extremely useful for integrating gender mainstreaming in a more complete, systematic and coherent way in any organization, including those involved in R&I.

2 Gender Mainstreaming as a Multifaceted Concept

Since the World Conference in Beijing in 1995 [9], it has not been controversial to state, both in the academic and institutional spheres, that equality between women and men can only be achieved if the gender perspective is mainstreamed right across the operation and performance of an organization, whatever its territorial level (international, regional, State, autonomous community or local) or its scope of competences (general or sectorial; such as R&I). Together with positive action, gender mainstreaming forms an essential strategy for making equality between women and men effective [10].

On the other hand, there is not the same level of consensus over the meaning that should be given to gender mainstreaming. This is an issue that, along with the theoretical implications it leaves unresolved, also has important practical consequences, as it directly affects the way it is integrated in organizations and in their ethical systems of governance [11–13].

If the available approaches are studied, it can be appreciated that both the way of understanding gender mainstreaming, and, above all, of making it operational, differ depending on the level at which it is overseen (global, regional, State or local); the degree of regulatory and institutional development of the equality of women and men (constitutional and legal framework, government structure, specialized agencies); the strength of the political commitment to equality; the budgetary effort devoted to public policies, and the technical and professional resources available to apply them, among many other factors [14, 15]. This assessment implies that the definition of gender mainstreaming is closely embedded with the political, institutional and organizational context in which it is supposed to take shape [16]. In fact, European R&I policies have sought a way of integrating the gender perspective, at least since the ETAN Report in 2001 [17].

If the most influential definitions available on mainstreaming are addressed [18], the Economic and Social Council of the United Nations, in its Agreed Conclusions (1997/2), sees it both as a *process* and as a *strategy*, in the following terms:[1]

> the process of assessing the implications for women and men of any planned action, including legislation, policies or programmes, in any field and at all levels. It is a strategy to make the experiences and concerns of women and men an integral part of the design, implementation, monitoring and evaluation of policies and programmes in all political, economic and social spheres, in such a way that women and men benefit equally, preventing the perpetuation of inequality. The ultimate goal is to achieve gender equality.

The Council of Europe, through its group of experts, defined gender mainstreaming in 1997, paying special attention to its *results* and to the *subjects* responsible for adopting it, as follows [19]:

> the (re)organisation, improvement, development and evaluation of policy processes, so that a gender equality perspective is incorporated in all policies at all levels and at all stages, by the actors normally involved in policy-making.

For its part, the European Union has understood gender mainstreaming as a *strategy* complementary to positive actions, in such a way that public intervention is not reduced to carrying out specific measures to favour women, but rather explicitly mobilizes all general actions and policies with a view to equality, actively and visibly taking into account in their design their possible effects on the respective situations of men and women ("gender perspective") [20]. For the *European Institute for Gender Equality* (EIGE), the European Union agency specializing in this issue[2], adopting gender mainstreaming involves integrating the gender perspective in the preparation, design, monitoring and evaluation of policies, regulatory measures and spending programmes, with the aim of promoting equality between women and men and fighting discrimination [21].

The Organization of American States has also addressed gender mainstreaming, which it has defined using the same terms as the United Nations Economic and Social Council [22] (Fig. 1).

Consideration of the meaning given to mainstreaming in the definitions compiled here allows us to see that it is conceived at the same time as a *process*, a *strategy*, an *evaluation* and some *results*. Above all, it does not have a specific enough meaning for it to be useful in determining what integrating it implies for a research institution [23] and for its development of responsible research and open science.

[1] Supplement No. 3 to the Official Records of the General Assembly (A/52/3/Rev.1), Chapter IV, paragraph 4.

[2] The *European Institute for Gender Equality* was established by Regulation 1922/2006, of the European Parliament and of the Council, of 20 December 2006 (Official Journal of the European Union L 403/9, of 30 December).

Body	Definition	Nature
United Nations (1997)	The process of assessing the implications for women and men of any planned action, including legislation, policies or programmes, in any field and at all levels.	Procedural Evaluator
Council of Europe (1999)	The (re)organization, improvement, development and evaluation of policy processes, so that a gender equality perspective is incorporated in all policies at all levels and at all stages.	Procedural Evaluator Obtaining results Transformative
European Union (1996, 2016)	Systematically taking into account the differences between the conditions, situations and needs of women and men in all Community policies and actions. This global and transversal approach requires an effort to mobilize all policies.	Analytical Systemic Approach
Organization of American States (2000)	The process of assessing the implications for women and men of any planned action, including legislation, policies or programmes, in any field and at all levels.	Procedural Evaluator

Fig. 1. Main international and regional definitions and identification of gender mainstreaming.

3 Gender Mainstreaming and R&I

The study of the gender issue, like the other key areas of RRI on the EU agenda (ethics, governance, public engagement, scientific education and open access), predates the formulation of this approach to research and innovation by the European Commission in 2010. While the literature on gender equality in R&I is abundant, with a history dating back at least to the 1990s [24], and while it continues to grow [25], the same is not true of its relationship with RRI either in the theoretical field of its conceptualization [26], or in the practice of RRI [27], or in the existing perceptions of RRI [28].

If, as most contributions do, we accept the European Commission's proposal [29], gender equality in R&I comprises three facets: *a) women's integration – horizontally as well as hierarchically – in all organizations working on R&I, b) integration of a gender perspective in policies and funding initiatives for the promotion of a structural change to identify implicit and explicit barriers, and c) integration of a gender perspective in research*. Intersectionality has recently been added as another facet of gender equality in R&I [30].

The literature that studies gender equality in R&I agrees on the persistence of gender barriers in academia and in research in the form of difficulties, obstacles and resistance that must be faced [31, 32]. The main limitations that the contributions have found are directly related to three of the facets that have been stated:

- the increase in the presence of women in all organizations that work in R&I [33], both horizontally and vertically (glass ceiling, sticky floor, slippery ladders and, particularly, leaky pipeline), due to both gender prejudices and stereotypes (gender bias) as well as job insecurity in science and innovation [34]. The imbalance is particularly clear in STEM areas of knowledge [35].
- The question of whether there are work-life balance policies or specific programs to improve the women's skills for promotion at work can also be an obstacle [24].
- the introduction of the gender perspective into actions or funding to produce a structural change [36].
- the integration of gender into the content and development of R&I and training [28].

The existence of many problems of gender equality implementation and evaluation in R&I and the general lack of a systematic understanding of the mechanisms underlying this issue have also been pointed out [24]; for example, having to permanently negotiate what gender equality means [31]. Some contributions point to a particular problem of outstanding importance for the purpose of this book: reducing gender equality to one of the key areas of RRI, instead of mainstreaming gender to integrate it into all dimensions of RRI [28].

While the facet that has most often been put into practice is the increase in women's participation in R&I, the aspect that has received the least attention is the introduction of gender as content and development of R&I [29]. One factor that explains this disparity is the autonomy of R&I organizations, especially university ones, in defining and implementing each of the four facets mentioned [34]. The idea is to carry out a *broad policy mix*, which individually supports the career of women in R&I while overcoming discriminatory institutional structures [37] – a mixed approach that has also been pointed out in other contributions [31].

There is no single mechanism to institutionalize the gender dimension in R&I, but different forms or models of institutionalization and governance have been identified, although not always sufficiently coordinated with one other:

- Developing and applying Gender Equality Plans in organizations and research institutions [38] or, in terms of organizational culture and governance, creating gender-friendly workplace cultures [39].
- Having a monitoring system that uses statistics and gender indicators [29].
- Establishing gender equality as a criterion for quality or excellence in research [28]; for example, introducing the gender perspective in the international mobility of research staff [40].
- Applying mechanisms for evaluating gender equality results [41].
- Enforcing national legislation on gender equality, approving charters or agreements with organizational or institutional principles, applying quotas, launching programmes or initiatives with specific funding – or modernizing the requirements for obtaining funding – providing information and supporting structures for small or local organizations [34]. Although the application of specific national legislation on gender equality institutionalizes this dimension, it also limits it, as R&I organizations limit themselves to complying with it without realizing they can also adopt other measures [27].

- Training in equality and identifying good practices in gender equality in the field of science and innovation [40].
- Training or skilling researchers in gender, without the need for them to become specialists in the field, has also been identified as a good practice [28]. An alternative is helping staff who are experts in gender equality with research, although this generates a degree of dependency [24].

A process of institutionalizing these instruments and mechanisms has been observed, although many of them were initially implemented as good practices (such as equality plans, monitoring, balanced presence, equality units, etc.). It has also been observed that the number and variety of governance tools available are directly related to the level of development of each of the four facets indicated above.

Sometimes, implementation is imposed by legal regulations in certain countries. In the case of the European Union, the *Gender Equality Plan* (GEP) has been chosen as the main tool for gender mainstreaming in R&I [42], as in 2022 it became an eligibility criterion for access to public and private institutions, research organizations and higher education bodies for funding from the *Horizon Europe Framework Program for Research and Innovation 2021–2027* [43].

In summary, the GEP must necessarily include two dimensions: firstly, four requirements related to the processes of the research organization: 1) its expression in a formal, public document, adopted by senior management and communicated within the institution; 2) the dedication of resources and expert knowledge for its implementation; 3) the collection of data broken down by gender and its monitoring through annual reports based on indicators; and 4) awareness-raising and training on gender equality; and, secondly, five recommended thematic areas: a) work-life balance and the organization's culture, b) the balanced presence of women and men in leadership and decision-making, c) equality in staff selection and professional promotion, d) the integration of the gender dimension in research and e) measures against gender violence including sexual harassment [44].

If the content of the GEP incorporates a good number of the mechanisms and instruments indicated throughout this section and makes it easier for R&I organizations to institutionalize them, the main result of its application should be that the gender mainstreaming cycle becomes standard practice in the organization. As a result there should be a structural change [45] that transforms the processes and results of the research it produces.

4 The European Institute for Gender Equality's Gender Mainstreaming Platform

Although GEPs represent a significant practical step forward in addressing the most pressing issues highlighted in the contributions on gender equality in R&I, there is still a long way together before gender mainstreaming is fully implemented in R&I organizations, at least if its content is compared with the contribution EIGE has made on gender mainstreaming with its *Gender Mainstreaming Platform* [21].

From a merely formal point of view, the EIGE Gender Mainstreaming Platform is a section of its website devoted to publicizing the way this EU agency understands gender

mainstreaming. However, from a material point of view, the EIGE's Gender Mainstreaming Platform simultaneously provides an Instruction Manual for mainstreaming the gender perspective and an Operation Manual so that gender mainstreaming can operate permanently. The EIGE Platform provides an *Instruction Manual* because this provides the components (or parts) of mainstreaming, describes them sufficiently, offers examples of good practices and, in many cases, provides manuals, guides or toolboxes to help implement them in organizations. In addition, the Platform also offers an *Operating Manual* for mainstreaming because, after having explained how these components or pieces should operate individually, it describes their joint and permanent operation through the *mainstreaming cycle.* That makes it interesting and useful for developing ethical governance systems for responsible research that mainstream gender equality (Fig. 2).

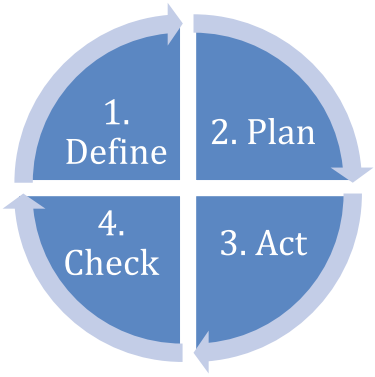


Fig. 2. Gender Mainstreaming Cycle.

According to the *mainstreaming cycle*, the gender perspective must be implemented in the definition, planning, implementation and evaluation processes of any organization, in both its internal or organizational dimension and in its external dimension (that intended for the public).

As can be seen, the Platform not only provides gender mainstreaming instruments or tools, it also places special emphasis on their interaction and on the importance of the processes to ensure their implementation is correct. Not surprisingly, methodology is one of the most characteristic features that the literature has highlighted in gender mainstreaming [11], to the point that it has been described as a concept-method [46].

The EIGE's Gender Mainstreaming Platform provides five components involved in the comprehensive application of this strategy in such a way that, without one or several of these five components, or with a limited or incomplete application of any of them, mainstreaming cannot be sufficiently integrated.

According to the EIGE, the five components of gender mainstreaming are the strategy, the dimensions, the conditions, the working methods and tools and, finally, the results. The five components mentioned, in turn, consist of a variable number of factors totalling 29 in all.

Figure 3 below shows the 29 factors of the Gender Mainstreaming Platform classified according to the five components they make up:

Component 1				
General mainstreaming strategy	Factors	1. Political commitment	2. Legal framework	
Component 2				
Dimensions	Factors	3. Balanced presence of women and men	4. Gender perspective in policy content	
Component 3				
Conditions	Factors	5. Implementation plan	8. Accountability mechanisms	11. Stakeholder participation
		6. Structures	9. Knowledge generation	
		7. Resources	10. Equality training	
Component 4				
Methods and tools	Factors	12. Gender analysis	18. Gender impact evaluation	24. Data broken down by gender
		13. Gender audit	19. Gender indicators	25. Consulting stakeholders on effective equality
		14. Awareness-raising	20. Monitoring	26. Institutional transformation
		15. Gender budgets	21. Planning	
		16. Equality training	22. Public procurement	
		17. Evaluation	23. Statistics	
Component 5				
Results	Factors	27. The institution's policies are better formulated	29. The institution's processes are more effective	
		28. The institution functions better		

Fig. 3. Components and factors of the EIGE's *Gender Mainstreaming Platform.*

The five components of the Platform, as well as the factors making up each component, are described in the following five subsections.

4.1 Component 1. The Gender Mainstreaming Strategy

The gender mainstreaming strategy involves two factors, without which it would no longer be a true mainstream strategy for effective equality of women and men: political commitment to mainstreaming and the appropriate legal framework for integrating and applying the strategy.

The R&I institution can promote the effective equality of women and men with different kinds of policies without these two factors (equal treatment, equal opportunities, even positive actions, etc.), or with only one of them, but both are essential for truly mainstream integration of the gender perspective.

The EIGE Platform also calls these two factors of the strategic component of mainstreaming "basic conditions", suggesting that they are fundamental for the strategy to be fully viable.

4.2 Component 2. The Two Dimensions of Gender Mainstreaming

Gender mainstreaming must be deployed in two dimensions: the presence of women and men in the organization, and the incorporation of equality between women and men as part of the content of its involvement in R&I. These are two factors that must once again be present in all the phases that the R&I intervention goes through (definition, planning, implementation and evaluation).

4.3 Component 3. The Conditions of Gender Mainstreaming

If legally recognizing gender mainstreaming as a strategy and publicly expressing commitment at the highest level to it are two "basic conditions" for integrating it across the board in an organization, they are not enough for mainstreaming to be deployed in the two dimensions we have mentioned: the balanced presence of women and men, and effective equality in the content of their R&I activities.

To mainstream the gender perspective, the institution also needs to meet at least seven conditions linked to its governance systems, at least to some extent: having a plan to implement gender mainstreaming; having organizational structures to carry it out;, having sufficient technical, human and economic resources to achieve it; being accountable during its implementation; learning more and better about gender mainstreaming during its implementation; having staff trained in gender mainstreaming; and having stakeholders involved in R&I.

Listing the factors in this order does not imply that they should be adopted consecutively or sequentially because, as has been noted, they are mutually interrelated, so the adoption of any one of them favours the implementation of others and, at the same time, all the factors suffer when one is not operational.

4.4 Component 4. Gender Mainstreaming Methods and Tools

The fourth component of the EIGE Gender Mainstreaming Platform contains the methods to be applied and the tools to be used to mainstream gender. This is a component that pays particular attention to the instrumental content of gender mainstreaming and is fundamentally operational. This component shows a particularly close relationship with *Component 3. Conditions of the Platform*, as it largely provides the necessary equipment for compliance with the conditions for gender mainstreaming.

The institution's work must use the following methods and tools so that the gender perspective is mainstreamed: gender analysis, gender auditing, awareness-raising, budgets with a gender perspective, equality training, evaluation, gender impact assessment, gender indicators, monitoring, planning, social equality clauses in public procurement, statistics, data broken down by gender, stakeholder consultation and institutional transformation.

It should be noted that the Platform does not precisely classify the 15 factors of this component into two groups, assigning certain factors to the set of methods and the rest to the set of tools. For example, it classifies *Factor 13. Gender audit* as a tool and as a method of mainstreaming; it describes *Factor 14. Awareness-raising* as a method and as a tool at the same time; it explains that *Factor 15. Budgets with a gender perspective* consists of both a tool and an approach to mainstreaming; and it considers that *Factor 16. Equality training* is not just another tool, but rather is part of a broader set of mainstreaming tools, instruments and strategies.

4.5 Component 5. The Results of Gender Mainstreaming

The fifth and final component ("Results") shows the results that must be derived from mainstreaming, formulated fundamentally in terms of improving the quality of the organization's intervention: policies are better formulated, the institution works better and its processes are more effective.

4.6 Characteristics of the Factors of the EIGE'S *Gender Mainstreaming Platform*

If the Platform's 29 factors are analysed, it is possible to list some characteristics that make it easier to understand them better, and also to draw up a classification in line with the character they share.

- The factors are very different, ranging from political commitment or the legal framework to measuring the improvements in the processes and the results the institution achieves, including the material resources it dedicates to this and the methods its staff uses. As a result they are not homogeneous, as they refer to a wide range of variables directly related to all facets of public intervention, from the binding legal framework to the training of staff responsible for their implementation.
- In many cases, the factors are mutually dependent, although with a different degree of intensity. Depending on their level of dependence or autonomy with respect to the rest of the factors, they can be classified into primary factors (compliance does not directly depend on other factors; for example, the regulatory framework of application, which comes before the dimensions, conditions, methods and results, and is not necessarily linked to political commitment) and secondary (compliance depends on the implementation of one or more of the other factors; for example, gender-sensitive evaluation, which depends on the gender analysis and evaluation as methodologies being adopted first).
- The way the implementation of each factor is demonstrated is not uniform either, as there are factors whose compliance can be shown through a single piece of evidence (simple factors; for example, an institutional declaration of political commitment to mainstreaming), while others need two or more pieces of evidence (complex factors; for example, awareness of equality).

Figure 4, inserted below, shows the classification of each factor depending on its degree of dependency on the other factors and according to the amount of evidence necessary to demonstrate compliance with it.

		Evidence of compliance	
		Simple	Complex
Interdependence with other factors	Dependent	2, 7, 10, 11, 14, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29	3, 4, 5, 6, 8, 9, 12, 13, 18
	Independent	1	

Fig. 4. Matrix of factors depending on their interdependence with other factors and the means of demonstrating compliance.

5 Discussion

Recognition of gender equality in the main international and European texts on human rights in the mid-20th century has encouraged the formulation of strategies and mechanisms to make it effective. This has been occurring in R&I at least since the beginning of the 21st century. Among these strategies, gender mainstreaming has occupied a prominent place for the last 30 years, even though it is a multifaceted concept whose operational application is neither direct nor easy.

Advances in gender equality and gender mainstreaming have been slow, costly and often fragile, and their development in R&I has not been any more straightforward. The reflection on gender equality in R&I is the accumulation of at least three decades of work, in which the difficulties, obstacles and resistance to achieving it have been sufficiently shown in relation to three facets: a) women's integration – horizontally as well as hierarchically – in all organizations working on R&I, b) integration of a gender perspective in policies and funding initiatives for the promotion of a structural change to identify implicit and explicit barriers, and c) integration of a gender perspective in research. Only recently has intersectionality been added as a fourth facet of the relationship between gender equality and R&I.

The reflection on gender equality and R&I has barely addressed the new approach represented by RRI. Two reasons can be put forward to explain this: it is an approach that, despite its potential, has not yet become widely established either in European culture nor in institutional practice on R&I; and RRI systematizes gender equality as one of its six keys without sufficiently noting that it is an overall dimension that must also be present in the other keys established by the European agenda: ethics, governance, public engagement, open access and scientific education.

Among the many gender mainstreaming methods and tools indicated to tackle the gender equality challenges in R&I, the Horizon Europe research funding program has chosen Gender Equality Plans (GEP) as a reference instrument for implementing gender mainstreaming in R&I organizations that want to obtain its aid. This decision represents progress for gender equality in R&I, as it addresses the main challenges in the facets that have been set out by institutionalizing a good number of gender mainstreaming methods

and tools. However, it does not exhaust all the possibilities available for an organization to fully integrate gender mainstreaming. Hence the importance and opportunity of mainstreaming gender equality in the ethical governance systems of research in R&I development and funding centres.

Through five components and 29 factors, the EIGE's Gender Mainstreaming Platform provides a more complete operational approach than the GEP for an R&I organization to integrate the gender mainstreaming cycle and achieve a true institutional transformation. However, the Platform maintains a complementary relationship with GEPs, as knowledge and the application of it can help appreciably not only in preparing, implementing and evaluating GEPs in better conditions, but also in enhancing their content to achieve better results in the processes, the operations and research results of R&I organizations.

The conclusions formulated allow us to argue that gender equality in RRI faces a double challenge: firstly, RRI itself is not yet definitively established in the European culture or institutional practice of R&I, at least if one takes into account the content of the Horizon Europe Framework Programme 2021–2027. And, secondly, there is the fact that the application of gender mainstreaming across the board, intended to produce a structural change in the organization, operation and activities of R&I organizations and, consequently, in the research and innovation stemming from them, is still far from being a standard in R&I governance.

Acknowledgements. This study was supported by the European Project "Ethics Governance System for RRI in Higher Education, Funding and Research Centres" [872360], funded by the Horizon 2020 programme of the European Commission.

References

1. Durán, P.: The pursuit of equality between women and men: the contribution of the United Nations. Cuadernos constitucionales de la Cátedra Fadrique Furió Ceriol **62–63**, 127–156 (2008)
2. Walby, S.: Gender Mainstreaming: productive tensions in theory and practice. Social politics. Int. Stud. Gend. State Soc. **12**(3), 321–343 (2005). https://doi.org/10.1093/sp/jxi018
3. Verloo, M.: Varieties of opposition to gender equality in Europe. Routledge (2018). https://doi.org/10.4324/9781315625744
4. Brisolara, S., Seigart, D., Sengupta, S.: Feminist Evaluation and Research: Theory and Practice. Guilford Press (2014)
5. Hubert, A., Stratigaki, M.: Twenty years of EU gender mainstreaming: rebirth out of the ashes? Femina Politica – Zeitschrift für feministische Politikwuissenschaft **25**(2), 21–36 (2016)
6. European Institute for Gender Equality. Gender Equality Index 2020: Digitalisation and the Future of Work. Publications Office of the European Union (2020). https://doi.org/10.2839/79077
7. European Institute for Gender Equality: Gender Equality Index 2021. Health. Publications Office of the European Union (2021). https://doi.org/10.2839/834132
8. European Commission, Directorate-General for Research and Innovation: Responsible Research and Innovation: Europe's Ability to Respond to Societal Challenges. Publications Office (2012). https://doi.org/10.2777/11739
9. Resolution 50/42, of 8 December 1995, of the United Nations General Assembly (A/CONF.177/20/Rev.1)
10. Durán, P.: The CEDAW's reception in Spain: the gap between law and practice. Yearb. Hum. Act. Hum. Rights **9**, 53–66 (2011)
11. Alonso, A.: El mainstreaming de género en España. Hacia un compromiso transversal con la igualdad. Tirant lo Blanch (2015)
12. Lombardo, E.: El mainstreaming: la aplicación de la transversalidad en la Unión Europea. Aequalitas: Revista Jurídica de igualdad de oportunidades entre mujeres y hombres **13**, 6–11 (2003)
13. Rey, M.: "Mainstream" de género. Eunomia. Revista en Cultura de la Legalidad **19**, 331–341 (2020)
14. Alfama, E., Cruells, M., De la Fuente, M.: ¿Qué ha cambiado con esta crisis? El mainstreaming de género en la encrucijada. Investigaciones feministas **5**, 69–95 (2014)
15. Bustelo, M., Lombardo, E.: Mainstreaming de género y análisis de los diferentes "marcos interpretativos" de las políticas de igualdad en Europa: el proyecto MAGEEQ. Aequalitas: Revista Jurídica de igualdad de oportunidades entre mujeres y hombres **17**, 15–26 (2005)
16. Alfama, E.: ¿Transformando el Estado? Avances y obstáculos en la implementación del mainstreaming de género. Ph.D thesis. Universidad Autónoma de Barcelona (2017). https://ddd.uab.cat/pub/tesis/2017/hdl_10803_458542/eaig1de1.pdf
17. European Commission: Research Directorate-General: Science policies in the European Union: Promoting excellence through mainstreaming gender equality. A Report from the ETAN Expert Working Group on Women and Science. Publications Office of the European Union (2001)
18. Durán, P.: Sobre el género y su tratamiento en las organizaciones internacionales. Ediciones Internacionales Universitarias (2007)
19. Council of Europe: Mainstreaming de género. Marco conceptual, metodología y presentación de "buenas prácticas". Final report of the activities of the group of experts on mainstreaming (EG-S-MS). Instituto de la Mujer (2005)

20. European Commission: Incorporating equal opportunities for women and men into all Community Policies and Activities. COM(96) 67 final (2012)
21. EIGE: What is gender mainstreaming? Publications Office of the European Union (2016). https://doi.org/10.2839/76981
22. OEA: Programa Interamericano sobre la promoción de los derechos humanos de la mujer y la equidad e igualdad de género. CIM/RES. 209/98 & AG/RES. 1732 (XXX-O/00) (2000)
23. Meier, P.: Implementing gender equality: gender mainstreaming or the gap between theory and practice. In: Hellsten, S.K., Holli, A.M., Daskalova, K. (eds.) Women's Citizenship and Political Rights. Palgrave Macmillan, New York (2006). https://doi.org/10.1057/9780230502901_10
24. Palmén, R., Kalpazidou Schmidt, E., Striebing, C., Reidl, S., Bührer, S., Groó, D.: Measuring gender in R&I – theories, methods, and experience. Interdiscip. Sci. Rev. **44**(2), 154–165 (2019). https://doi.org/10.1080/03080188.2019.1603873
25. European Institute for Gender Equality: Gender in Research. Publications Office of the European Union (2022). https://doi.org/10.2839/062112
26. Burget, M., Bardone, E., Pedaste, M.: Definitions and conceptual dimensions of responsible research and innovation: a literature review. Sci. Eng. Ethics **23**(1), 1–19 (2016). https://doi.org/10.1007/s11948-016-9782-1
27. Stahl, B.C., Chatfield, K., Ten Holter, C., Brem, A.: Ethics in corporate research and development: can responsible research and innovation approaches aid sustainability? J. Clean. Prod. **239**, 1–14 (2019). https://doi.org/10.1016/j.jclepro.2019.118044
28. Bührer, S., Wroblweski, A.: The practice and perceptions of RRI – a gender perspective. Eval. Program Plann. **77**, 1 (2019). https://doi.org/10.1016/j.evalprogplan.2019.101717
29. European Commission, Directorate-General for Research and Innovation: Indicators for promoting and monitoring Responsible Research and Innovation. Report on the Expert Group on Policy Indicators for Responsible Research and Innovation. Publications Office of the European Union (2015). https://doi.org/10.2777/9742
30. European Commission: A new ERA for Research and Innovation. COM (2020) 628 final (2020)
31. Cacace, M., Balahur, D., Bleijenbergh, I., Falcinelli, D., Friedrich, M., Schmidt, E.K.: Structural transformation to achieve gender equality in science guidelines. Stages (2013). https://eige.europa.eu/sites/default/files/guidelines_stages_4.pdf
32. European Commission, Directorate-General for Research and Innovation: She figures 2021: gender in research and innovation: statistics and indicators. Publications Office (2021). https://doi.org/10.2777/06090
33. Wroblewski, A., Bührer, S., Leitner, A., Fan, C.: Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI). Analytical Report on the Gender Equality Dimension. European Commission (2015)
34. Lipinsky, A.: Gender Equality Policies in Public Research. Publications Office of the European Union (2014). https://doi.org/10.2777/65956
35. European Commission, Directorate-General for Research and Innovation: Bridging the Gender Gap in STEM: Strengthening Opportunities for Women in Research and Innovation. Publications Office of the European Union (2022). https://data.europa.eu/doi/10.2777/774922
36. Espinosa, J., Bustelo, M., Velasco, M.: Evaluating gender structural change. guidelines for evaluating gender equality action plans. University of Bradford (2016). https://bit.ly/3zIlAkq
37. Meijer, I., Mejlgaard, N., Woolley, R., Rafols, I., Wroblewski, A.: Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) – a Preliminary Framework for RRI Dimensions & Indicators (2016)
38. ERAC SWG on Gender in Research and Innovation: Gender Equality Plans as a Catalyst for Change. ERAC 1202/21 (2021)

39. Wojniak, J.: Responsible research and innovation. How to put gender equality into practice? Zarządzanie Publiczne **2**(38), 163–176 (2017). https://doi.org/10.4467/20843968ZP.16.014.7229
40. Otero-Hermida, P., García-Melón, M.: Gender equality indicators for research and innovation from a responsible perspective: the case of Spain. Sustainability **10**, 2980 (2018). https://doi.org/10.3390/su10092980
41. Espinosa-Fajardo, J., Bustelo, M., Velasco M.: ¿Cómo evaluar el cambio estructural de género en universidades? Aprendizajes de una práctica colaborativa y feminista. IgualdadES **7**, 413–440 (2022). https://doi.org/10.18042/cepc/IgdES.7.03
42. European Commission, Directorate-General for Research and Innovation: Horizon Europe Guidance on Gender Equality Plans (GEP). Publications Office of the European Union (2021). https://data.europa.eu/doi/10.2777/876509
43. Art. 7.6 of Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013 (Text with EEA relevance)
44. European Institute for Gender Equality: Gender Equality in Academia and Research. GEAR Tool Step-by-step Guide. Publications Office of the European Union (2022). https://doi.org/10.2839/354799
45. European Commission, Directorate-General for Research and Innovation: Horizon Europe Guidance on Gender Equality Plans (GEP). Publications Office of the European Union (2021). https://data.europa.eu/doi/10.2777/083753
46. Dauphin, S. and Sénac-Slawinski, R.: Gender mainstreaming: analyses des enjeux d'un "concept-méthode" (2008). https://hal-sciencespo.archives-ouvertes.fr/hal-01559371v2/document

Navigating the Future and Overcoming Challenges to Unlock Open Science

Eva Méndez(✉) and Pablo Sánchez-Núñez

OpenScienceLab, Department of Library and Information Science, Universidad Carlos III de Madrid, Getafe, Spain
{emendez,psnunez}@bib.uc3m.es

Abstract. Open Science (OS) has been rapidly evolving in recent years, but there is still work to be done to return Science to researchers and citizens who pay for it. Technological advancements have enabled Open Science to transform the way scientific research is conducted, facilitating collaboration and innovation among researchers. As a result, OS is expected to play an increasingly important role in scientific research and innovation in the years to come, driving discoveries and advancements in various fields. However, OS also poses challenges, including the potential for bias and discrimination in research. This chapter explores the challenges that need to be addressed to fully implement OS globally, outlining the barriers that need to be overcome and describing the complexity of the changes that come with this new research approach. Additionally, the chapter discusses the impact of Artificial Intelligence on addressing these challenges, while also creating new ones.

Keywords: Open Science · Science Policy · Scholarly Communication · Research Evaluation · Intellectual Property · Artificial Intelligence · Responsible Research and Innovation (RRI)

1 Introduction: Open Science in RRI

Responsible Research and Innovation (RRI) implies a philosophical and far-reaching ethical approach, which assesses society's expectations regarding research and innovation, to foster equity and sustainability in research and innovation. As other chapters of this book explain, RRI includes Open Access (OA) as an extra/added/common element, along with ethical issues, gender equality, citizen participation or public engagement, scientific education, and governance of the scientific and innovative process. Stahl defined RRI as a "higher-level responsibility or meta-responsibility that aims to shape, maintain, develop, coordinate and align existing and novel research and innovation-related processes, actors, and responsibilities to ensure desirable and acceptable research outcomes" [1]. RRI's ethical principles and meta-responsibility include open access to knowledge, but also entail all the processes and challenges of Open Science (OS). The RRI concept has been used to describe a new way of governing research and the relationships between the agents involved, called the quadruple helix when we talk about innovation: academia,

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 203–223, 2023.
https://doi.org/10.1007/978-3-031-33177-0_13

industry, government and civil society [2]. These four players are also the key stakeholders in making Open Science a reality. The orchestration of these four stakeholders and how their relationship and engagement in OS is articulated is key to overcoming OS's current challenges.

Research and Innovation are becoming more and more digital, complex, data-driven and so reliant on powerful computing capabilities, which has given rise to e-Science and scientific computing. This leads to a better understanding of complex scientific problems, faster progress in scientific discovery, and more accurate and reliable scientific information. Meanwhile, digital technologies, particularly the World-Wide Web, which facilitates "distributed collaborative research behaviour" [3] and the possibility of immediately, openly and massively communicating knowledge through the network, lead us to think about the promise of the transformation of science and the opening of research processes.

Helped along by these trends, Open Science has been gaining significant relevance in the past ten years, and it is expected to play a crucial role in shaping the future of scientific research and innovation [3]. Despite that, we finance and carry out research in the same way as in the last century and publish and assess it as we began doing several centuries ago. Institutions, researchers, and founders are hostages to an anachronistic, absurd and ineffective scientific communication model based exclusively on 'the scientific paper' as an end in itself. This 'paper-centric' model often leads to a simplistic conception of Open Science as 'Open Access to scientific publications', just as it simplifies the quality of research to a high Journal Impact Factor (JIF). However, the last SWAFS Program report [4] includes OS rather than Open Access as a fundamental element of RRI, highlighting the importance of the fuller meaning of Open Science.

There are different systematic reviews around the definition of Open Science [5, 6], and a plethora of terms to characterise or explain its essence: a movement, a trend, a paradigm, a construct, or an attitude [7], in the scientific research ecosystem. All the definitions emphasise transparency, collaboration, and access to data, publications, methodologies, software and any other research outcome. OS aims to promote scientific knowledge that is accessible to all, rather than being restricted to a select few. It has become a key element of Responsible Research and Innovation that promotes scientific advance by enabling researchers to collaborate, co-create and build upon each other's work, fostering innovation by making the latest information and developments easily accessible, increasing public trust in science by promoting reproducibility through more transparent scientific data and methods, and supporting interdisciplinary research by encouraging collaboration between different fields. All this conceptualization of Open Science is right, but our favourite approach is to explain Open Science as giving Science back to the researchers that do it and to the citizens that pay for it. Figure 1 represents the particular view of Eva Méndez of OS's key components. Responsible Research and Innovation is the ethical and philosophical foundation that allows us to believe in Open Science. But the roots, which should cement the OS system, are research infrastructures, scientific integrity and a new research evaluation system that can incentivise Open Science's practices.

Open Science (OS) is an achievable goal due to technical and policy advances. But if the research cycle cannot be shared, data cannot be reused, citizens are not involved

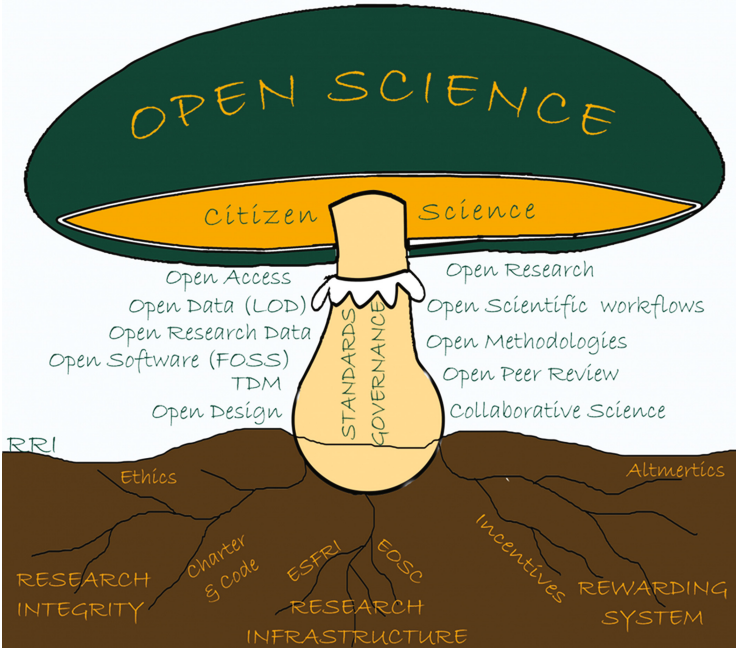


Fig. 1. Open Science Mushroom done by Judit Eva's Fazekas-Paragh (OpenAire NOAD (https://www.openaire.eu/blogs/hungary-on-the-move-1)) based on Eva Méndez's Open Science Mushroom [7].

and researchers cannot be assessed by their whole contribution to the system and OS practices, and the goals of OS will remain aspirations beyond the practical reach of many. Since we started to speak about OS as the new paradigm to make better, more efficient and transparent science, we constantly talk about 'drivers and barriers' [8], claiming to pass to action[1], but always pointing out at a lot of challenges [10–13].

This chapter shares the personal views of the authors on their meta-research for Open Science, based on their involvement in different groups and fora (e.g., OSPP[2], RDA[3], CoARA[4], YERUN Open Scienge WG[5], etc.) to evolve OS and make it a reality. The chapter offers a perspective on the current status of Open Science in the light of

[1] This was the sprit of the "Amsterdam Call For Action" in 2016 under the Dutch presidency. Source: https://www.openaccess.nl/sites/www.openaccess.nl/files/documenten/amsterdam-call-for-action-on-open-science.pdf and also the work of the OSPP (Open Science Policy Platform) during the second mandate where it was defined the PCIs concept: Practical Commitments for Implementation of Open Science[9].

[2] https://openscience.eu/open-science-policy-platform-final-report

[3] https://www.rd-alliance.org

[4] Https://coara.eu

[5] https://yerun.eu/work/open-science

European policies and initiatives[6], UNESCO Recommendation[7], and the urgent changes needed to make Open Science the default approach. Those key changes are: the way we communicate Science; the way we assess research, researchers, and research performing organisations (RPOs); and the way we give credit in a new legal framework that protects research results, without leaving anybody behind. It will not be easy, but by addressing these challenges, the scientific community should work together to create a more stimulating, transparent, accessible, fair, and equitable research ecosystem. Finally, we discuss the opportunity and challenge of Artificial Intelligence (AI) for Open Science implementation and the Responsible Research and Innovation ecosystem.

2 Framing up the 'Right to Science' and a Global Science Commons

Open Science (OS) is based on the principle that scientific knowledge should be freely available to everyone, without restrictions [6]. This means that scientific research should be conducted openly and transparently, with results being made publicly available through open-access publications and open-data platforms. Also, OS is based on the belief that it can promote greater collaboration, innovation, and scientific progress, and can help to address global challenges such as poverty, inequality, and environmental degradation.

The OS movement, in the "democratic school" approach, emphasises that access to knowledge is a human right, so the social return on public investments in science (accountability) is the guarantee of this universal right [14]. The *Universal Declaration of Human Rights*[8] (1948, art. 27) established the fundamental right to science and culture. It is also recognized in other international declarations underlining the importance of science in the promotion of development and ensuring that individuals can "enjoy the benefits of scientific progress and its applications" (*International Covenant on Economic, Social and Cultural Rights*[9] 1966, art. 15). However, it is often treated as a privilege. The International Covenant also recognizes the "freedom of the researcher" as an indispensable value, as well as cooperation in science[10]. Even though OS is a newer term, its values and principles are more than 50 years old. The essence of science is openness and collaboration, understanding it from the Mertonian perspective of knowledge as a 'common' or 'public good' defended among others by Hess and Ostrom [12, 15–17]. Geiger and Jütte have furthermore underlined the importance of "a right to research" substantiated in the European fundamental rights instruments (freedom of expression, freedom of arts and sciences and right to education) and defend a new definition of a 'right to research' to remove copyright barriers in favour of research [18].

[6] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en

[7] https://en.unesco.org/science-sustainable-future/open-science/recommendation

[8] https://www.un.org/en/about-us/universal-declaration-of-human-rights

[9] https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights

[10] *Ibid.*

The United Nations has been supporting the discussion on a Global Science Commons since 2019 to discuss the key role of OS in the achievement of the UN 2030 Agenda, outlining a Science Commons as the framework organised around principles, universal values and the architecture of open research, and based in OS as a key accelerator of the Sustainable Development Goals[11]. In 2021, at the UN's meeting on Open Science, the "right to science[12]" was highlighted, and this year (2023), the meeting is focused on the idea of a *Global Open Science Commons for the Sustainable Development Goals,* based on three main topics: equity in open scholarship, reforming scientific publishing, and strengthening the science-policy-society interface[13].

The discussion is centred around the key principle that researchers have the freedom to perform research, and that citizens have the right to access research outputs. The universal 'right to science', the 'right to research' and the principles of OS are closely linked, as they aim to promote greater access to scientific knowledge and to ensure that the benefits of scientific progress are shared by everyone.

UNESCO has played a key role in promoting the principles of OS and advocating for greater access to scientific knowledge and information. The *UNESCO Recommendation on Science and Scientific Researchers* [19] was adopted by the UNESCO General Conference in 2017. This first recommendation does not speak about "Open Science" directly (it refers to open access, data, educational resources, software, etc.), but it sets out several principles and guidelines for promoting science and scientific research that will ground the very recommendation of Open Science [20], including the importance of freedom of scientific expression and inquiry, the promotion of ethical and responsible scientific practices, and the need to ensure that the benefits of scientific advances are shared fairly and equitably. Since the adoption of the UNESCO recommendation on science and scientific research, there has been significant growth and development in the field of OS, with new initiatives, and practices emerging to support open and collaborative scientific research (e.g. Hong Kong Principles[14], CoARA[15]).

The *Recommendation on Open Science* [20] was endorsed by the 193 members of the UN in November 2021. It provides a framework for promoting OS principles, including transparency, collaboration, and access to scientific information. The recommendation aims to promote the democratisation of science, reduce barriers to access and participation, and ensure that scientific knowledge is used to benefit all societies (Fig. 2).

This recommendation aims to provide a/the ultimate "vital tool" to improve the availability and quality of both scientific outcomes and scientific processes to bridge the gaps in science, technology, and innovation among different countries, in order to fulfil the human right of access to science. It offers governments, research institutions, and the scientific community a set of precepts and principles for promoting and implementing OS. It recognizes the importance of OS for advancing scientific knowledge, improving the quality of research, and enhancing the impact of science on society. The

[11] https://research.un.org/ld.php?content_id=51390330 (November 2019).

[12] https://www.un.org/sites/un2.un.org/files/open_science_outcome_document_v.3b.pdf (November 2021).

[13] https://www.un.org/en/library/OS23 (February 2023).

[14] https://www.wcrif.org/hong-kong-principles

[15] Https://coara.eu

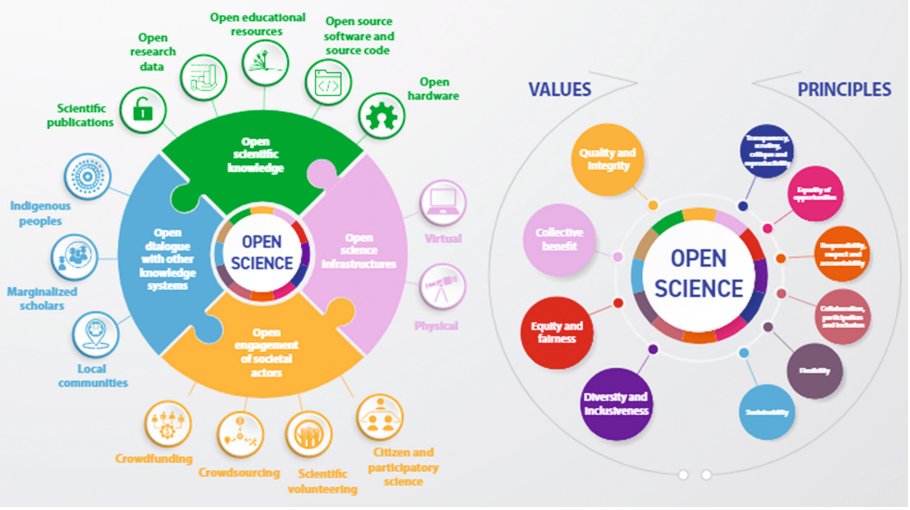


Fig. 2. Understanding Open Science (UNESCO Open Science Toolkit[16]).

recommendation also acknowledges the challenges posed by the rapid pace of technological change, and it calls for the development of international norms and standards to ensure that OS practices are consistently applied across different regions and scientific disciplines [21].

UNESCO is "the legitimate global organisation enabled to build a coherent vision of open science and a shared set of overarching principles and shared values[17]", but "it is the response of working scientists, their institutions, and funders that will determine whether a new mode of Open Science is achieved as the 'new normal' and whether it realises the hopes of its proponents" [22]. The UNESCO document is not 'yet another declaration', it has developed an implementation plan with working groups and a steering committee[18], as well as a toolkit with guides to put the recommendations into practice[19]. The toolkit can benefit institutions in understanding Open Science, developing policies, funding OS, building capacity or bolstering OS infrastructures for all. The Open Scholarship Initiative (OSI), working in partnership with UNESCO, published a comprehensive roadmap for Open Science in light of the Recommendation, calling for a collective future for OS and Open Research. However, the future of Open Science is in the hands of individual researchers and the institutions underpinning their work, while the target of UNESCO recommendations is the member countries, which have different economic situations and different limitations in their funding, practices and procedures.

16 https://unesdoc.unesco.org/ark:/48223/pf0000383323

17 https://council.science/current/news/the-questionnaire-unesco-open-science-recommendation/

18 https://www.unesco.org/en/open-science/implementation

19 https://www.unesco.org/en/open-science/toolkit

3 Global Challenges for Open Science: Rocky Pathways Ahead

Open Science is a tremendously diverse and interconnected space. "Reforming it will not be as simple as claiming that open is X, the solution is Y, and the path to the future can be enforced by a unilaterally-developed mandate" [23]. Despite its potential benefits, the future of OS is surrounded by big challenges, real changes and new frameworks [10], as well as the engagement of many stakeholders.

From the very beginning, the European Commission (EC) started to define the elements of OS as 8 key challenges: four related to research outcomes (FAIR data, the European Open Science Cloud (EOSC)), next-generation metrics, and new scholarly communication mechanisms); and four challenges related to research stakeholders (evaluation of researchers' careers, skills, and training in OS, new research integrity, and citizen science)[20]. The term 'challenges', like "barriers and drivers", has always been used in the spirit of putting OS to work.

Several authors started to think about the future of Open Science and to point out at the challenges, realities, and current limitations [15, 27] even for publishers [28]. Pownall et al. [13] envision some key trends and developments for the future of OS, noting:

- Technological advancements make it easier for researchers to share and access data and software**.** The widespread adoption of cloud computing, for example, has made it possible for researchers to store, process, and analyse massive amounts of data from various sources. This, in turn, has facilitated collaboration between researchers, as they can now share data, publications, etc. and conduct 'real-time research' [28]. These developments are expected to make it easier for researchers to access and use data, as well as to collaborate on software development in different fields [29], but also in cross-disciplinary research.
- Artificial intelligence (AI) is seen as a particularly powerful tool for scientific research [30]. These technologies have the potential to revolutionise the way scientific research is conducted, by allowing researchers to process and analyse vast amounts of data more efficiently. For example, AI can be used to analyse complex biological data, such as genomic data, and identify new drug targets for diseases. But AI also deals with ethical issues, as pointed out by ETHNA project [31]. AI entails new developments to analyse the impact of OS [32] better and faster (Cf. Section 4 below.).
- Open Science is also expected to play a critical role in addressing societal challenges such as climate change, healthcare, and poverty, among others.
- Traditionally, scientific research has been highly fragmented, with researchers in different fields often working in isolation. However, Open Science has the potential to break down these barriers and facilitate collaboration between researchers from different fields, by advancing interdisciplinary research and enabling them to tackle complex scientific problems that cannot be solved by a single discipline, for example,

[20] All these EU challenges have been addressed by the Open Science Policy Platform (OSPP), the EU High-Level Advisory Group run by the European Commission from 2016–2020. See. [24] and also by different institutions, like LERU, YERUN, EUA, etc. See for example the document targeting Universities and Research Performing Organizations [25].

to address Sustainable Development Goals [33] or an extreme crisis like COVID-19 [34, 35]. The rights to research and to access research outcomes can help guarantee sustainability, innovation, and justice [18].

There are many studies, surveys and different approaches to identifying OS's challenges [10, 36, 37]. These challenges include issues related to data privacy and security, funding, data quality, data compatibility, data ownership and control, expertise in data management and analysis, etc. Other approaches focus on the challenges of real practice of OS, on the understanding, practicality, transparency, sharing, and replication from particular disciplines like biosciences or psychology [26].

Beyond new data-driven research approaches or the practice of OS inside a particular discipline, real global challenges come from the current traditional 'research business'. A paradigm shift is needed [7, 22], even a "profound shift in epistemology" [38]. Here we detail four challenges to OS becoming a default paradigm for new research: three urgent changes, plus the move to a real and fair framework of diversity, equity, inclusion, and accessibility principles.

3.1 Challenge #1: Change the Way We Communicate Science

Traditionally, scholarly communication has been mediated by academic journals and publishers, which act as gatekeepers for the publication of scientific research. However, this system has been criticised for being expensive, slow, dysfunctional, and market-based, primarily serving the purpose of bringing profit to commercial publishers[21].

The oligopoly of a few scientific publishers has become the most profitable current legal business, with annual revenues comparable to Google, Apple, or Amazon[22]. This leaves researchers dependent on the "most profitable obsolete technology in history"[23]. This approach is not good for scientific communication, but by restricting access to the scientific record it also risks losing the public's trust in science, undercutting global inclusion, and largely failing to realise the opportunities presented by the digital revolution. It seems that at the moment, misinformation and disinformation on scientific advances are freely available online to all, while credible and authoritative scientific information and data lie behind paywalls.

Scholarly communication faces many challenges, including information overload, limited access to scientific knowledge, and quality issues like questionable research practices [39] or research reproducibility and integrity [40], among problems frequently discussed [41]. Innovative approaches to scholarly communication are needed to ensure that

[21] See the explanatory article published in *The Guardian* by Stephen Buranyi, in 2017 describing the scientific publishing system. Is the staggeringly profitable business of scientific publishing bad for science?. Available in: https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science

[22] See. https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science

[23] Academic journals: The most profitable obsolete technology in history. This was the title of a post by Jason Schmitt in the Huffington Post Blog, on December 2014. Available in: https://www.huffpost.com/entry/academic-journals-the-mos_b_6368204

scientific research is shared and communicated effectively. Traditionally, the scientific evaluation system forced researchers to focus on writing academic papers. Publishing a 'paper' has become an objective in itself, and researchers seem to be just 'paper-makers' instead of solving the world's scientific challenges.

The current Open Science/Open Access approaches maintain the 'paper-centric' model while embracing open-access publishing, renewing the business of publishers by charging for publishing (Article Processing Charges or APCs), or even worse, charging for reading as well as publishing (transformative agreements[24]), and generating other issues like the so-called 'predatory journals'[42]. Another approach is to adopt new technologies, such as preprint servers, repositories, etc. allowing researchers to quickly and easily share their findings with the scientific community before formal (paid) publication (preprints) or after the publication is issued in a journal with an embargo period (repositories). Different open access policies (Fig. 3), described by colours (gold, green, diamond, even 'black' considering rogue solutions such as Sci-Hub)[25], widely used, but they have not yet individually or collectively ensured all papers are openly available [43], nor do they seem likely to do so.

Paying for publishing —instead of paying for reading— only brings forward the payment, and a uniform cost is added to all research funding. The more predictable revenue stream no longer relies on anyone prioritising the money to buy journals. It primarily creates more profit for the big publishers, who charge for hybrid journals (the ones where researchers are encouraged to publish if they want promotion or funding). PlanS, cOAlitionS[26] and R&P (read and publish) transformative deals have unintentionally reinforced the power and the business of large publishers. As Science Business pointed out recently: "The big problem with open access publishing is that it's expensive. Science is a big business, with more than $2 trillion spent on R&D each year, according to UNESCO, making the business of publishing research results a big cash cow"[27]. Solutions like diamond open access publishing platforms (e.g. Open Research Europe[28]) are a better option since they make research papers freely available to both readers and authors, so scholarly communication structures (articles/publications/publishers) are more sustainable and inclusive in the long term [44]. But retaining the centrality of the traditional paper also maintains the opportunity window for traditional academic publishers to

[24] Transformative agreements aim to transition subscription-based publishing to open access, but issues arise around the costs and sustainability of these agreements, as well as concerns about the impact on smaller publishers and potential conflicts of interest. For a better understanding of transformative agreements, see this blogpost from Pandelis Pearakakis: https://pandelisperakakis.info/2021/05/07/what-are-transformative-agreements-and-what-to-know-before-using-them/

[25] See Lucy Barnes' blogpost to review all the OA routes, including hybrid, bronze, libre and gratis: https://blogs.openbookpublishers.com/green-gold-diamond-black-what-does-it-all-mean

[26] https://www.coalition-s.org

[27] Science Business sumarized the EU council conclusions under the Swedish presidency https://sciencebusiness.net/news/Universities/leaked-eu-member-states-set-out-reform-scientific-publishing For the draft of the Council conclusions see: https://data.consilium.europa.eu/doc/document/ST-5997-2023-INIT/en/pdf

[28] https://open-research-europe.ec.europa.eu

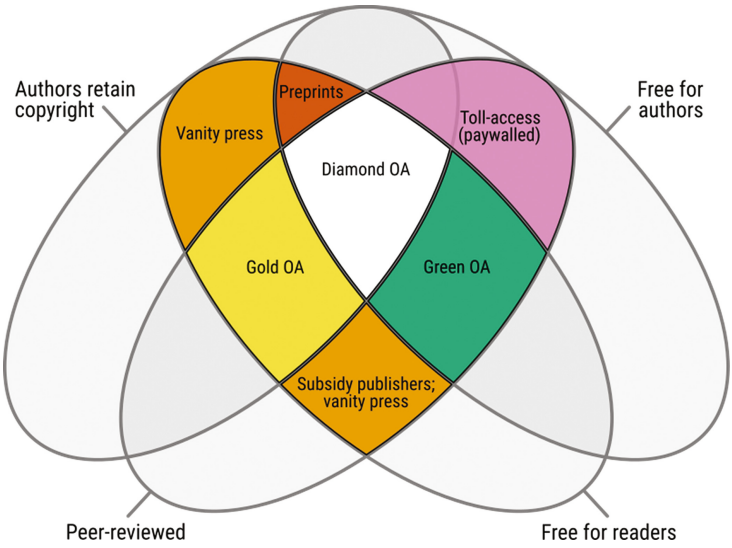


Fig. 3. Venn diagram highlighting the different levels of open access in scholarly publishing, as a function of cost to the readers and authors, copyright retention, and peer review (Farquharson, Jamie (2022): Diamond open access venn diagram [in SVG]. Figshare. Figure: https://doi.org/10.6084/m9.figshare.21598179.v1).

continue to reap handsome profits while holding back scientific progress through their outdated business models, whose effects have now become regressive.

A disruptive solution is (very much) needed, coherent with current technology and needs. Researchers need to cut out the whole notion of publishing, "get rid of papers[29]", pages, and journals, and return as much as possible of the outrageous costs that this outdated business extracts as rents to fund the work of research. While research papers have long been the cornerstone of scientific communication, and they made sense 350 years ago —when the first scientific journal was created[30]— in the current context of digital transformation and new data-driven research, they are senseless and inefficient.

The very notion of a "paper" should be completely overhauled. Research papers are often lengthy and difficult to understand. As a format, they are readily prone to errors that are not easy to detect[31]. This makes it difficult for researchers to keep up with the latest developments in their field, and for citizens to understand the significance of new findings. We have reached a point where, perhaps with supervision, Artificial Intelligence can write papers for us. The most important outcome of research is not a final article, but the hypothesis, the demonstrations, the data, and the methodologies,

[29] The Guardian reflected again in April 2022 on the issue of scientific publishing by Stuart Ritchie, The big idea: should we get rid of the scientific paper?Available: https://www.theguardian.com/books/2022/apr/11/the-big-idea-should-we-get-rid-of-the-scientific-paper

[30] https://royalsociety.org/journals/publishing-activities/publishing350/history-philosophical-transactions/

[31] See: Ball, P. It's not just you: science papers are getting harder to read. *Nature* (2017). https://doi.org/10.1038/nature.2017.21751

often including software to analyse them. These factors open up the scientific cycle and enable collaboration with researchers from different disciplines who can advance science more efficiently than at present, even if they do not normally follow academic journals in those areas.

Policymakers prioritize what they believe is best for scientific research as a whole, but they may not take into account the needs and concerns of individual researchers 'all for the researcher, but without the researcher[32]. It is essential to ask actual researchers what they want, and how they can improve their research communication, without those proposals being provided on the researchers' behalf by commercial publishers. Recently, the Open Scholarship Initiative published the results of their "Research Communication Survey", where 90% of respondents selected some version of the answer "I think there are better ways of doing research communication" and 75% wanted to hear about and explore new ideas and policies [45].

3.2 Challenge #2: Change the Way We Assess Research and Researchers' Evaluation

This has always been the hottest core issue under the debate of what we need to truly implement Open Science: "to change the way we measure science[33]". We need a complete reset of the system for how research is valued, measured and assessed.

Along with the anachronistic scholarly communication system described above, contemporary research assessment is outdated and inappropriate for currently needed research. It provides existential support for the old-fashioned and dysfunctional paper-centric system, primarily enabling the unsustainable and disproportionately profitable business around it.

Traditionally, scientific research's quality and impact are evaluated using metrics such as the journal impact factor (JIF). JIF measures the frequency with which the average article in a given journal has been cited in a particular year. "JIF in policy and decision-making in academia is based on false beliefs and unwarranted inferences" [46]. But we cannot blame the system for using JIF in the evaluation procedures of researchers' careers and grants, unless we provide them with an alternative measure or mechanism. Quantitative metrics (JIF, h-index, etc.) have been largely criticised for being too narrow and for not accurately reflecting the true impact of scientific research [47–51]. Many important contributions to science may not receive numerous citations, while studies with questionable methodology or data may receive a high number of citations simply because they were published in prestigious journals.

Alternative ways to assess a researcher's contribution to Science are also discussed: altmetrics, new indicators, next-generation metrics, etc. [52–55], but without a clear consensual approach. One potential approach would be to focus on alternative metrics including the so-called altmetrics [53] and/or social media metrics [56] such as the number of downloads and views of a research article, or the number of times a dataset has

[32] See: Eva Méndez (@evamen) comments to Commissioner Moedas' speech on European Open Science Cloud (EOSC) Summit (June 2017):https://yerun.eu/wp-content/uploads/2017/07/Dr-Eva-Mendez-input-to-EOSC-Summit.pdf

[33] Watch the statement in the Introduction video Conference Open Science 4–5 April 2016, Dutch Presidency EU2016: https://www.youtube.com/watch?v=C9a3Ap3yyak&t=277s

been reused in other research studies. But none of these alternative metrics (also not immune to certain problems or limitations) has been adopted consistently in the evaluation procedures. The only advance so far, is the particular and growing commitment of some institutions to adhere to the DORA declaration[34], Leiden manifesto[35], or more recently, the Hong Kong principles[36] and SCOPE principles[37]. All these declarations and principles pledge to reward more qualitative factors (sharing practices, peer review, students' advice and mentoring, etc.) and ground the evaluation on the researcher instead of the outputs —such as sole publications. These new approaches to embracing more qualitative assessments of researchers and research careers acknowledge the values of OS in establishing criteria beyond purely quantitative indicators, like the CAM (Career Assessment Matrix) defined by the European Commission expert group [49]. This new trend tries to give room to everyone's talent in academia[38], as well as Responsible Research Assessment (RRA) centred on the researcher [57]. RRA[39] does not have a single universally agreed definition, but it "is an umbrella term for approaches to assessment which incentivise, reflect and reward the plural characteristics of high-quality research, in support of diverse and inclusive research cultures" [58]. The recent launch of the Coalition for Advancing Research Assessment (CoARA[40]) marks a significant step toward the responsible use of metrics, and qualitative and comprehensive research evaluation. CoARA is a bottom-up initiative that resulted from the European Commission's facilitation to propose new approaches to research evaluation [59]. CoARA will seek and discuss alternative mechanisms to evaluate scientific research, beyond traditional metrics, such as JIF, and focus on the actual impact and reach of research papers. This new approach places a greater emphasis on open-access publishing, data sharing, and collaboration, reflecting the growing recognition that these elements are critical components of high-quality scientific research. At the time of writing, more than 450 institutions have signed the agreement[41] and more than 410 have also joined the coalition. By adhering to CoARA, the scientific community can actively participate, pilot and move on to new research assessment. Scientists need this 'coalition of doers' but the technology to come up with real alternative research metrics is also needed. In this sense, the recently approved project GraspOS[42] (Next Generation Research Assessment to Promote Open Science) might come up with a technological solution that becomes

[34] San Francisco Declaration on Research Assessment (DORA). 2012: https://sfdora.org

[35] Leiden Manifesto for Research Metrics. 2015: http://www.leidenmanifesto.org

[36] Hong Kong principles for assessing researchers. World Conference on Research Integrity. 2019.https://www.wcrif.org/guidance/hong-kong-principles

[37] SCOPE Framework for Research Evaluation (iNORMS). 2021.https://inorms.net/scope-framework-for-research-evaluation

[38] See the position paper by the Universities in the Netherlands. 2020. Room for everyone's talent: towards a new balance in the recognition and rewards of academics: https://www.universiteitenvannederland.nl/recognitionandrewards/wp-content/uploads/2019/11/Position-paper-Room-for-everyone%E2%80%99s-talent.pdf

[39] https://www.leidenmadtrics.nl/articles/navigating-responsible-research-assessment-guidelines

[40] Https://coara.eu

[41] https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf

[42] https://cordis.europa.eu/project/id/101095129

an alternative to the current JIFs hiding also under paywalls in commercial data sources like Web of Science (WoS) and Scopus.

Research assessment is a complex area to change, and the approach must be multifaceted. It should be shaped by collaboration with researchers. It requires a new assessment equation that includes all the talents and merits that researchers and scientists have. Likewise, it should provide transparency and accountability, enable the assessment of interdisciplinary and cross-disciplinary research and its inclusiveness and equity. In addition, it is essential the impact of research on society and the environment, and recognising the increasing importance of OS it should promote better and broader collaboration. Research evaluation needs to be dynamic, flexible, and innovative, leveraging the latest technology and trends to ensure that the best and most impactful research is being funded and recognised.

3.3 Challenge #3: Change the Way We Give Credit to Scientific Contributions. A Step Forward Copyright

Intellectual Property Rights (IPR), such as patents and copyrights, play a critical traditional role in protecting the rights of researchers, publishers and/or innovators. These rights were developed to give researchers control over the use and dissemination of their work and provide a mechanism for them to be compensated for their contributions to the scientific enterprise. However, they can also become a barrier to OS, as they may limit the ability of others to build upon and use existing knowledge. This can lead to decreased collaboration and information sharing, reducing public access to scientific research in turn.

While IPR has long been seen as a critical component of scientific innovation, there is a growing recognition of the fact that current approaches can be a barrier to OS [60]. The current EU legislative framework (the Information Society Directive[43], Database Directive[44] and the Digital Single Market Directive[45]) contains relevant provisions for access to and reuse of scientific publications and data. In the current legislative landscape, researchers or their institutions are the first copyright owners of exclusive rights over the scientific publications they produce. However, publication in "high-quality" journals —which today is generally strongly linked to career development in academia— very often requires transferring those rights to publishers or giving them exclusive licences. The exclusive rights of copyright owners contrast with the fragmentation of exceptions and limitations to such rights. The current framework also poses challenges for the protection of researchers as authors and their interest in fostering the dissemination of knowledge. Although copyright rules were drafted to protect authors' creations, the widespread practice of researchers assigning copyright to publishers does not result in favouring the interests of those originally intended to benefit.

There is growing recognition that a balance must be struck between protecting IPR and promoting OS [61, 62]. There are several approaches to move forward with the re-use of scientific outcomes (not only publications but data). A traditional option is to adopt

[43] https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=CELEX%3A02001L0029-20190606
[44] https://eur-lex.europa.eu/eli/dir/1996/9/2019-06-06
[45] https://eur-lex.europa.eu/eli/dir/2019/790/oj

more open recognition of moral rights with open licences, such as the Creative Commons licences, which allow researchers to share and build upon existing knowledge while still protecting their rights. In Europe, there is a specific action (2) in the European Research Area to "propose an EU copyright and data legislative framework for research[46]". There are legislative proposals to introduce a "secondary publication right" for publicly funded scientific publications, and non-legislative ones like those developed by funders (e.g. the Rights Retention Strategy, introduced by cOAlition S[47]) or by institutions (e.g. the Harvard model*, or the #ZeroEmbargo campaign run by LIBER[48]).

The future of IPR itself holds changes and challenges. As well as confronting the increasing importance of digital technology, and the growing importance of artificial intelligence, IPR policies are being reconsidered in light of the growing recognition of the value of OS and the concomitant need to balance the interests of creators and users, as well as the need to address issues related to globalisation including access and affordability. Current IPR standards and policies need to keep pace with technological developments, especially online, to avoid legal uncertainty negatively affecting knowledge production in research. It is urgent to address new copyright and IPR regimes to guarantee better intellectual protection that can support open, transparent and collaborative science.

3.4 Challenge #4: EDI-A (Equity, Diversity and Inclusion-Accessibility) in Open Science

Equity, Diversity, and Inclusion (or EDI) are increasingly important for research-performing organisations, and its principles are becoming crucial in Open Science, particularly since the UNESCO recommendation [20]. Embracing OS principles, researchers, and research institutions can help to build a more equitable and fair world. These recommendations and principles are well accepted on paper, but the reality is that we are far away from them being consistently and effectively applied in practice. Accessibility is one of the FAIR (Findable, Accessible, Interoperable and Reusable) principles applied to research data-sharing practices, and in general, to any research outcomes including software [63]. However, accessibility in this context is used to mean the availability of research outcomes. There is another significant sense of the term, commonly thought of as "web accessibility" due to the substantial and widely recognised work of W3C in this area*. This is critical to ensure the rights of people with disabilities to carry out research and to have equitable access to scientific knowledge.

Sustainable Development Goal 10 focuses on reducing inequalities and promoting social, economic and political inclusion for all[49]. Participation in OS is not occurring on an even playing field. The structural inequalities that shape our societies also manifest within academia, particularly in terms of unequal access to resources, and therefore, some operate within the Open Science space at a distinct advantage relative to others [64]. Current 'naive solutions' like "replacing big subscription deals with big APC deals simply flips inequity in accessing content with inequity in publishing content" [43].

[46] https://era.gv.at/public/documents/4678/02_Copyright_and_data_legilsative_framework.docx
[47] https://www.coalition-s.org/rights-retention-strategy
[48] Https://libereurope.eu/zeroembargo
[49] https://sdgs.un.org/goals/goal10

Failing to address structural inequalities directly, means that the advantages of those who are already privileged will grow, especially as they have a greater influence on how OS is implemented [65]. A policy framework with practical implementation that ensures Equity, Diversity, and Inclusion (and access to people with disabilities) is needed if we are to achieve the Sustainable Development Goal 10 with regard to accessing scientific advances.

4 Open Science and Artificial Intelligence: A Necessary Synergy

Open Science (OS) aims to make scientific knowledge more accessible and transparent. Artificial intelligence (AI) promises to bring new insights and capabilities to scientific research [66]. OS and AI are evolving at the same time, but we require them to evolve working together for the benefit of Science in the digital transformation era[50].

4.1 Combining OS and AI

We need to harness synergetic advances in OS and AI, to support the coming Open Scientific paradigm. However, there are a number of challenges in doing this in a scientific research scenario:

- Combining OS and AI can be technically challenging, since it requires researchers to manage and analyse large amounts of data using complex AI algorithms. This can require specialised training and resources, and it is difficult for researchers to keep up with the rapid pace of technological change in the field [67].
- OS relies on making scientific data publicly available, which can raise concerns about data privacy and security. AI can be used to analyse this data, but it must be anonymized to protect individuals' privacy. Ensuring data privacy and security helps to build trust in OS practices and protects against potential ethical and legal violations. This is especially relevant in areas like medical, economic and sociological research, where sensitive personal information is often involved [69].
- AI can perpetuate and amplify biases in scientific research. This typically happens when the overall data used to train AI algorithms is biased in some way. This is a significant concern in OS, where data is made publicly available and can be used by a wide number of researchers [68].
- The traditional 'paper-centric' model of scientific publishing, with a heavy emphasis on research papers, is giving way to more innovative and accessible forms of scientific communication. AI algorithms and tools like ChatGPT [66, 70] are playing an increasingly important role in scientific communication, for example being listed as an author on research papers [6]. There is growing recognition that data itself is becoming the most valuable currency of scientific research. "Data is essential for AI, as algorithms in these systems need large quantities of high-quality data to perform properly and to develop further by 'learning'. The continuous promotion of data quality within

[50] The European Commission is aware of this need on its priorities for funding the European Open Science Cloud (EOSC) https://cordis.europa.eu/project/id/101058593 https://ai4eosc.eu/ See. https://ai4eosc.eu/

organisations when using (open and FAIR) data for AI is therefore essential to gain reliable insights[51]". By embracing OS, and moving to place a greater emphasis on sharing research outcomes as FAIR data than on the production of papers, scientists can continue to advance our understanding of the world and improve the quality of scientific research.

4.2 New Scenarios for Research Integrity

Technology is not neutral. The use of AI in scientific research presents many opportunities, but also several challenges, particularly concerning research integrity (See also Chapter 10, in this volume): the principles of honesty, trustworthiness, and transparency that underpin the scientific process. In the context of AI, these principles must be upheld, especially as AI systems become increasingly complex, and their reasoning becomes difficult to reproduce. Hence, the urgent need for a culture of Responsible AI [71].

Examples of how research integrity can be compromised by AI in scientific research are (1) bias in training data (if the training data used to build an AI system contains biases, many AI systems perpetuate and amplify these biases in their results); (2) lack of transparency (many AI systems, particularly deep learning models, can be difficult to interpret, making it difficult for researchers to understand why the AI system is making certain decisions); (3) overreliance on AI results (as AI systems become increasingly sophisticated, there is a risk that researchers become too reliant on the results produced by these systems, even if they cannot understand how the AI system reached these results); and (4) misuse of AI systems. There is a risk that AI systems may be used in ways that violate ethical principles or compromise research integrity [72].

5 Conclusions: Open Science for Today (and Tomorrow)

The old-fashioned practices (journals, papers, impact factors, h-index etc.) are designed to create and maintain the 'publish or perish' status quo that drives researchers' careers and academic publishers' profits all over the world. The gold open-access approach will leave a lot of researchers, countries and poorly funded disciplines behind. It often seems that a researcher's fundamental goal is 'to write papers', rather than understanding and finding solutions for global issues. By embracing OS and finding innovative approaches to scholarly communication, the scientific community can continue to advance our understanding of the world and improve the quality of scientific research. In this chapter, we have revisited the main changes necessary for a research paradigm built around the practice of Open Science.

The traditional and current publication and evaluation system was conceived for research in a qualitatively different era, and it is not suited to contemporary and emerging needs of the scientific "system". A new research evaluation system must evolve to reflect changes in research practices and emerging technologies, ensuring that it remains relevant and effective. We can not simply replace the JIF with another (more updated,

[51] AI and Open Data: a crucial combination (data.europa.eu, 2018): https://data.europa.eu/en/publications/datastories/ai-and-open-data-crucial-combination

but equally flawed) metric or indicator. We require a conclusive system that, applying all the current technologies (AI, blockchain, etc.) can come up with an updated and efficient mechanism to evaluate current science, in a twofold sense: assessing all that is needed to be evaluated in the current research performance and applying all the current technologies.

Digital transformation is about changing the entire system using web-era technologies and principles. Every change entails an effort, and the real epistemological shift might be uncomfortable for incumbents (publishers, but also a proportion of academics who are comfortably positioned to benefit from the status quo). To promote open access and redesign the current system, simply switching from printed journals to digital PDFs is not enough. Business models must be revised to better support open access, and outdated metrics should be reevaluated with a sense of responsibility. Additionally, IPR legislation should include more exceptions to better align with these changes. In the interest of an equitable society, equipped to develop and apply scientific knowledge to the pursuit of its goals, we need fundamental changes in the whole ecosystem that determines the way we do and communicate research.

Of course, there are many other challenges ahead to make Open Science a holistic reality (citizen science, FAIR data in the data research infrastructures like EOSC or the research integrity), but the ones detailed here are urgent changes needed to firmly fulfil OS goals.

Acknowledgement. The authors want to thank Charles Nevile for English proofreading and his valuable suggestions for the final version of this document.

References

1. Stahl, B.C.: Responsible research and innovation: The role of privacy in an emerging framework. Sci. Public Policy **40**, 708–716 (2013). https://doi.org/10.1093/scipol/sct067
2. McAdam, M., Debackere, K.: Beyond 'triple helix' toward 'quadruple helix' models in regional innovation systems: implications for theory and practice: Beyond 'triple helix' toward 'quadruple helix' models. RD Manag. **48**, 3–6 (2018). https://doi.org/10.1111/radm.12309
3. David, P.A., Besten, M.D., Schroeder, R.: How Open is e-Science? In: 2006 Second IEEE International Conference on e-Science and Grid Computing (e-Science 2006), pp. 33–33. IEEE (2006). https://doi.org/10.1109/E-SCIENCE.2006.261117
4. Delaney, N., et al.: Science with and for society in Horizon 2020: achievements and recommendations for Horizon Europe. In: European CommissionDirectorate-General for Research and Innovation (2020). https://doi.org/10.2777/32018
5. Stracke, C.M.: Open science and radical solutions for diversity, equity and quality in research: a literature review of different research schools, philosophies and frameworks and their potential impact on science and education. In: Burgos, D. (ed.) Radical Solutions and Open Science. LNET, pp. 17–37. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-4276-3_2
6. Vicente-Saez, R., Martinez-Fuentes, C.: Open science now: a systematic literature review for an integrated definition. J. Bus. Res. **88**, 428–436 (2018). https://doi.org/10.1016/j.jbusres.2017.12.043

7. Méndez, E.: Open science por defecto. La nueva normalidad para la investigación. Arbor **197**, a587 (2021). https://doi.org/10.3989/arbor.2021.799002
8. Hessels, L.K., Koens, L., Diederen, P.J.M.: Perspectives on the future of open science: effects of global variation in open science practices on the European research system. In: Publications Office (2021). https://doi.org/10.2777/054281
9. Méndez, E., Rebeca, L., MacCallum, C.J., Moar, E., Open Science Policy Platform Members.: Towards a shared research knowledge system : final report of the open science policy platform. In: European Commission. Directorate-General for Research and Innovation (2020). https://doi.org/10.2777/00139
10. Allen, C., Mehler, D.M.A.: Open science challenges, benefits and tips in early career and beyond. PLOS Biol. **17**, e3000246 (2019). https://doi.org/10.1371/journal.pbio.3000246
11. Altman, M., Cohen, P.N.: The scholarly knowledge ecosystem: challenges and opportunities for the field of information. Front. Res. Metr. Anal. **6**, 751553 (2022). https://doi.org/10.3389/frma.2021.751553
12. Krishna, V.V.: Open science and its enemies: challenges for a sustainable science-society social contract. J. Open Innov. Technol. Mark. Complex. **6**, 61 (2020). https://doi.org/10.3390/joitmc6030061
13. Pownall, M., Talbot, C.V., Kilby, L., Branney, P.: Opportunities, challenges and tensions: Open science through a lens of qualitative social psychology. Br. J. Soc. Psychol. bjso.12628 (2023). https://doi.org/10.1111/bjso.12628
14. Bartling, S., Friesike, S.: Towards another scientific revolution. In: Bartling, S., Friesike, S. (eds.) Opening Science, pp. 3–15. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-00026-8_1
15. Boyle, J.: Mertonianism Unbound?: Imagining free, decentralized access to most cultural and scientific material. In: Ostrom, E., Hess, C. (eds.) Understanding Knowledge As a Commons: From Theory to Practice. The MIT Press (2007)
16. Hensher, M., Kish, K., Farley, J., Quilley, S., Zywert, K.: Open knowledge commons versus privatized gain in a fractured information ecology: lessons from COVID-19 for the future of sustainability. Glob. Sustain. **3**, e26 (2020). https://doi.org/10.1017/sus.2020.21
17. Understanding knowledge as a commons: from Theory to Practice. The MIT Press (2007). https://doi.org/10.7551/mitpress/6980.001.0001
18. Geiger, C., Jütte, B.J.: The right to research as guarantor for sustainability, innovation and justice in EU copyright law. In: Taina, E., Pihlajarinne, J.M., Pratyush U. (eds.) Intellectual Property Rights in the Post Pandemic World: An Integrated Framework of Sustainability, Innovation and Global Justice. Edward Elgar Publishing (2022)
19. UNESCO. Recommendation on Science and Scientific Researchers (ANNEX II), pp. 117–145 https://unesdoc.unesco.org/ark:/48223/pf0000260889.page=116 (2018)
20. UNESCO. UNESCO Recommendation on Open Science. https://unesdoc.unesco.org/ark:/48223/pf0000379949 (2021). https://doi.org/10.54677/MNMH8546
21. Bronner, M., Meijer, G., Yam, V., Friedrich, B.: UNESCO issues a powerful endorsement of Open Science. Nat. Sci. **2**, (2022). https://doi.org/10.1002/ntls.10037
22. Cudennec, C., Sud, M., Boulton, G.: Governing open science. Hydrol. Sci. J. 1–4 (2022). https://doi.org/10.1080/02626667.2022.2086462
23. Hampson, G., et al.: Open Science Roadmap. Open Scholarsh. Initiat. Proc. **2020** (2020). https://doi.org/10.13021/OSI2020.2735
24. OSPP-REC. OSPP-REC open science policy platform recommendations. In: European Commission. Directorate-General for Research and Innovation (2018). https://doi.org/10.2777/958647
25. Ayris, P., Maes, K., López de San Román, A., Labastida, I.: Open Science and its role in universities: A roadmap for cultural change. (2018).

26. Guzzo, R.A., Schneider, B., Nalbantian, H.R.: Open science, closed doors: the perils and potential of open science for research in practice. Ind. Organ. Psychol. **15**, 495–515 (2022). https://doi.org/10.1017/iop.2022.61.
27. Naaman, K., et al.: Exploring enablers and barriers to implementing the Transparency and Openness Promotion Guidelines: a theory-based survey of journal editors. R. Soc. Open Sci. **10**, 221093 (2023). https://doi.org/10.1098/rsos.221093
28. Hocquet, A.: open science in times of coronavirus: introducing the concept of 'Real-Time' Publication. Substantia 937 (2020). https://doi.org/10.13128/SUBSTANTIA-937
29. Scheffler, M., et al.: FAIR data enabling new horizons for materials research. Nature **604**, 635–642 (2022). https://doi.org/10.1038/s41586-022-04501-x
30. Xu, Y., et al.: Artificial intelligence: a powerful paradigm for scientific research. The Innovation **2**, 100179 (2021). https://doi.org/10.1016/j.xinn.2021.100179
31. González-Esteban y Patrici Calvo, E.: Ethically governing artificial intelligence in the field of scientific research and innovation. Heliyon **8**, e08946 (2022). https://doi.org/10.1016/j.heliyon.2022.e08946
32. Wang, K.: Opportunities in open science with AI. Front. Big Data **2**, 26 (2019). https://doi.org/10.3389/fdata.2019.00026
33. Giovannini, E., et al.: The role of science, technology and innovation policies to foster the implementation of the sustainable development goals (SDGs): report of the expert group "Follow up to Rio+20, notably the SDGs". Publications Office (2015). https://doi.org/10.2777/615177
34. Besançon, L., et al.: Open science saves lives: lessons from the COVID-19 pandemic. (2020). https://doi.org/10.1101/2020.08.13.249847
35. Homolak, J., Kodvanj, I., Virag, D.: Preliminary Analysis of COVID-19 Academic Information Patterns: A Call for Open Science in the Times of Closed Borders. https://www.preprints.org/manuscript/202003.0443/v1 (2020). https://doi.org/10.20944/preprints202003.0443.v1
36. Heise, C., Pearce, J.M.: From open access to open science: the path from scientific reality to open scientific communication. SAGE Open **10**, 215824402091590 (2020). https://doi.org/10.1177/2158244020915900
37. Scheliga, K., Friesike, S.: Putting open science into practice: a social dilemma? First Monday (2014). https://doi.org/10.5210/fm.v19i9.5381
38. Mirowski, P.: The future(s) of open science. Soc. Stud. Sci. **48**, 171–203 (2018). https://doi.org/10.1177/0306312718772086
39. Spector, P.E.: Is open science rewarding A while hoping for B? Ind. Organ. Psychol. **15**, 516–519 (2022). https://doi.org/10.1017/iop.2022.64
40. Fanelli, D.: Is science really facing a reproducibility crisis, and do we need it to? Proc. Natl. Acad. Sci. **115**, 2628–2631 (2018). https://doi.org/10.1073/pnas.1708272114
41. Tennant, J.P., et al.: Ten hot topics around scholarly publishing. Publications **7**, 34 (2019). https://doi.org/10.3390/publications7020034
42. Grudniewicz, A., et al.: Predatory journals: no definition, no defence. Nature **576**, 210–212 (2019). https://doi.org/10.1038/d41586-019-03759-y
43. Green, T.: Is open access affordable? Why current models do not work and why we need internet-era transformation of scholarly communications. Learn. Publ. **32**, 13–25 (2019). https://doi.org/10.1002/leap.1219
44. Becerril, A., et al.: OA Diamond Journals Study. Part 2: Recommendations. https://zenodo.org/record/4562790 (2021). https://doi.org/10.5281/ZENODO.4562790
45. Hampson, G.: Summary of OSI2022 Research Communication Surveys. forthcoming (2023)
46. Paulus, F.M., Cruz, N., Krach, S.: The Impact Factor Fallacy. Front. Psychol. **9**, 1487 (2018). https://doi.org/10.3389/fpsyg.2018.01487

47. Curry, S., Gadd, E., Wilsdon, J.: Harnessing the Metric Tide: indicators, infrastructures & priorities for UK responsible research assessment. 11014215 Bytes https://rori.figshare.com/articles/report/Harnessing_the_Metric_Tide/21701624 (2022). https://doi.org/10.6084/M9.FIGSHARE.21701624
48. Hatch, A., Curry, S.: Changing how we evaluate research is difficult, but not impossible. eLife **9**, e58654 (2020). https://doi.org/10.7554/eLife.58654
49. O'Carroll, C., et al.: Evaluation of research careers fully acknowledging Open Science practices: rewards, incentives and/or recognition for researchers practicing Open Science. http://dx.publications.europa.eu/10.2777/75255 (2017). https://doi.org/10.2777/75255
50. de Rijcke, S., Wouters, P.F., Rushforth, A.D., Franssen, T.P., Hammarfelt, B.: Evaluation practices and effects of indicator use—a literature review. Res. Eval. **25**, 161–169 (2016). https://doi.org/10.1093/reseval/rvv038
51. Thomas, D.A., Nedeva, M., Tirado, M.M., Jacob, M.: Changing research on research evaluation: a critical literature review to revisit the agenda. Res. Eval. **29**, 275–288 (2020). https://doi.org/10.1093/reseval/rvaa008
52. Fenner, M.: Altmetrics and other novel measures for scientific impact. In: Bartling, S., Friesike, S. (eds.) Opening Science, pp. 179–189. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-00026-8_12
53. Torres-Salinas, D., Cabezas-Clavijo, Á., Jiménez-Contreras, E.: Altmetrics: new indicators for scientific communication in Web 2.0. Comunicar **21**, 53–60 (2013). https://doi.org/10.3916/C41-2013-05
54. Wilsdon, J., et al.: Next-generation metrics: responsible metrics and evaluation for open science (2017). https://doi.org/10.2777/337729
55. Wouters, P., et al.: Indicator frameworks for fostering open knowledge practices in science and scholarship (2019). https://doi.org/10.2777/445286
56. Wouters, P., Zahedi, Z., Costas, R.: Social media metrics for new research evaluation. In: Glänzel, W., Moed, H.F., Schmoch, U., Thelwall, M. (eds.) Springer Handbook of Science and Technology Indicators. SH, pp. 687–713. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-02511-3_26
57. Susi, T., et al.: Centrality of researchers in reforming research assessment Routes to improve research by aligning rewards with Open Science practices (2022)
58. Stephen, C., et al.: The changing role of funders in responsible research assessment: progress, obstacles and the way ahead (RoRI Working Paper No.3). 2449096 Bytes https://rori.figshare.com/articles/report/The_changing_role_of_funders_in_responsible_research_assessment_progress_obstacles_and_the_way_ahead/13227914 (2022). https://doi.org/10.6084/M9.FIGSHARE.13227914
59. European Commission. Directorate General for Research and Innovation. Towards a reform of the research assessment system: scoping report. European Commission. Publications Office (2021). https://doi.org/10.2777/707440
60. Rhoten, D., Powell, W.W.: The frontiers of intellectual property: expanded protection versus new models of open science. Annu. Rev. Law Soc. Sci. **3**, 345–373 (2007). https://doi.org/10.1146/annurev.lawsocsci.3.081806.112900
61. de la Cueva, J., Méndez, E.: Open science and intellectual property rights: How can they better interact? : State of the art and reflections : executive summary. Publications Office (2022). https://doi.org/10.2777/347305
62. Willinsky, J.: Copyright's Broken Promise: How to Restore the Law's Ability to Promote the Progress of Science. Massachusetts Institute of Technology, Cambridge (2023)
63. Chue Hong, N., et al.: FAIR Principles for Research Software (FAIR4RS Principles) (2021). https://doi.org/10.15497/RDA00065

64. Cole, N.L., Reichmann, S., Ross-Hellauer, T.: Toward equitable open research: stakeholder co-created recommendations for research institutions, funders and researchers. R. Soc. Open Sci. **10**, 221460 (2023). https://doi.org/10.1098/rsos.221460
65. Ross-Hellauer, T.: Open science, done wrong, will compound inequities. Nature **603**, 363 (2022). https://doi.org/10.1038/d41586-022-00724-0
66. Stokel-Walker, C., Van Noorden, R.: What ChatGPT and generative AI mean for science. Nature **614**, 214–216 (2023). https://doi.org/10.1038/d41586-023-00340-6
67. Thesmar, D., et al.: Combining the power of artificial intelligence with the richness of healthcare claims data: opportunities and challenges. PharmacoEconomics **37**, 745–752 (2019). https://doi.org/10.1007/s40273-019-00777-6
68. Ntoutsi, E., et al.: Bias in data driven artificial intelligence systems—An introductory survey. WIREs Data Min. Knowl. Discov. **10** (2020). https://doi.org/10.1002/widm.1356
69. Naudé, W.: Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. AI Soc. **35**(3), 761–765 (2020). https://doi.org/10.1007/s00146-020-00978-0
70. Stokel-Walker, C.: ChatGPT listed as author on research papers: many scientists disapprove. Nature **613**, 620–621 (2023). https://doi.org/10.1038/d41586-023-00107-z
71. Dignum, V.: Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Springer International Publishing, Cham (2019)
72. Gundersen, O.E., Gil, Y., Aha, D.W.: On reproducible AI: towards reproducible research, open science, and digital scholarship in AI publications. AI Mag. **39**, 56–68 (2018). https://doi.org/10.1609/aimag.v39i3.2816

Ethics and Development of Advanced Technology Systems in Public Administration

António da Costa Alexandre[1] and Luís Moniz Pereira[2(✉)]

[1] Lisbon, Portugal
antoniocalexndre@outlook.pt
[2] Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Monte da Caparica, Lisbon, Portugal
lmp@fct.unl.pt

Abstract. This article analyses aspects related to the implementation of advanced technology systems in Public Administration, taking into account the scope of action of its entities and services, as well as its main distinction with the private sector, which lies in the pursuit of the public interest, understood as justification for the execution of public policies, but also in the broader scope of the human development index. The difficulties and resistance to the adoption and use of technology in Public Administration will be examined, also in the light of practical situations whose implementation proved to be inadequate, resulting in the analysis of ideas for the future. We approach the evolution of public management models that are emerging through the action of technology. We prepare the ground for machine ethics in Public Administration by framing ethics in public services in general terms, with reference to AI systems designed in line with mainstream ethics. Of the various ethical issues that arise in this domain, we pay attention to the issue of privacy and the balance that needs to be achieved so that the use of data can contribute to ethical, beneficial and reliable technologies. We sound alerts in the field of discrimination and prejudice that the bias of technology can show. The explainability and transparency of technological systems provide confidence to decision-makers and citizens, helping to clarify responsibilities in the decisions of Public Administration agents, and the researchers working for them .

Keywords: Ethics · Advanced Technology · Public Administration

1 Introduction

In this article we list some Public Administration (PA) definitions that allow delimiting the field of action of public services and entities [1, 2], which, due to the breadth of areas in which they operate, can be the main buyer of technology, without neglecting the regulatory functions that the PA is responsible for ensuring. The theoretical framework is complemented by data from interviews carried out in other academic investigations.

A. da Costa Alexandre—Independent Researcher.

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, pp. 224–247, 2023.
https://doi.org/10.1007/978-3-031-33177-0_14

The public interest is the main distinguishing factor of PA vis-à-vis the private sector, whose main objective is to achieve economic benefits, which can justify the search for non-profit technologies, in favour of benefits and improvement of the human development index.

We frame the adoption of technology by public bodies and services by studying two well reported known cases: both regarding the implementation of fraud risk assessment systems in granting social support, in the Netherlands and in the United States [3]. Although with operational distinctions, beneficiaries are penalised by having subsidies refused in spite of the absence of grounds, and by the inability of the systems to provide the reasons for the refusals, so that the latter cannot be contested.

Technology is also causing an evolution in public management models of PA entities and services, with the emergence of new paradigms, closer to citizens through the technological systems that assist them in carrying out tasks and implementing public policies [4, 5].

The machine ethics approach will be preceded by the theoretical framework of ethics in PA, as a necessary condition to prepare an environment conducive to ethical principles that indicate solutions to questions and problems that may arise from the implementation of technology in general.

The process of conceiving, adopting and using technologies is analysed, in an attempt to identify facilitating conditions in terms of the research and organisational culture, leadership, plus the advantages that technology can bring to public services and entities in terms of effectiveness, efficiency, transparency, and trust of citizens.

Privacy is one of the key points for citizens to trust the technological solutions that support PA bodies, without neglecting the need to find a balance between data protection and necessary concessions that allow progress, facilitate security, or contribute to public health [6].

Discrimination caused by technology arising from technological solutions that give continuity to prejudices present in society, due to the use of non-representative, inaccurate or simply wrong training data, is also mentioned [7]. Examples are presented of systems for evaluating the risk of recidivism in criminal activity used by USA judges in the decision to grant parole, which penalise citizens of African origin.

The question of the explainability of technological solutions is analysed from a dual perspective, in terms of the information that citizens have the right to obtain on issues they raise before the PA, but also in terms of attributing responsibilities to service agents and public entities [8].

2 Public Administration, Public Interest and Human Development

The expression "Public Administration" (PA) can be understood according to two different meanings: (i) regarding the mode of organisation with the characteristics and specificities of the institutions that comprise it, with a view to achieving the public interest, in compliance with the law and rights, and (ii) regarding organisational action, object of study in the sociology of organisations [1].

Several versions of the PA definition were compiled by [9], with the aim of listing the PA concepts that many conceived to demonstrate the scope and nature of the discipline

of PA. The perception of the concept of PA is important, from the outset, to reach the meaning of the word "administration" as an effort that requires cooperation from the intervening parties in the pursuit of a common objective. The PA participates in the organisation of public policies and government programs [9].

According to [2] it is easy to define PA if one is satisfied with a simplistic definition. In that case, it will be the government in action managing public affairs and implementing public policies, in a public context—without which the PA does not exist—, plus affirming its role in the execution of governmental activity. These authors highlight the decisive role of the PA in the implementation of the public interest, defining it as the "universal label" that involves public policies and their various execution programs. The public interest is generally considered an asset by the community. However, less simplistically, it can be used both to promote public policies that represent an effective common good, and to obscure other policies whose acceptance by the community is not so evident. Walter Lippmann, cited by [2] affirmed "the public interest may be presumed to be what humans would choose if they saw clearly, thought rationally, and acted disinterestedly and benevolently" [10].

Haeberlin and Comim [11] refer that [12] considered that the replacement of religious beliefs, as a principle of government, with "belief in the public interest" was at the origin of the construction of modern societies. These authors defend the insufficiency of a legal approach *stricto sensu* of public interest, advocating a broader approach that contemplates human development.

Sen [13] in his studies on ethics and economics, argues that the criterion of utility should be replaced by the notion of substantive freedoms, considering the autonomy of free choice and the real conditions of a life that makes sense to value it. This author influenced the first United Nations document on human development.

The Human Development Index (HDI) has been established since 1990, viz. in [14]) issued by the United Nations Development Program (UNDP). In the various developments it has undergone, it takes on the role of contributing to correction of distortions in the model proposed by the HDI, in what concerns the role of poverty and of inequalities in development.

According to [15] "human development can provide the possibility of longer, healthier, and more creative lives, achieving goals that humans value, with commitments that shape development with a view to equality and sustainability of the planet we share."

3 Technologies and Public Administration

Digitization processes tend to replace existing processes by altering state governance and labour relations, intensifying development tensions, as a result of the growing trend to implement Artificial Intelligence (AI) [16].

Giest and Klievink [3] understand that because of digitization and the increasing use of systems based on algorithms and data processing, a new service provision regime is being created. Technologies such as AI alter technical fields, communication channels, decision-making functions and mechanisms, as well as their levels of intervention control [17] and the ethics thereof.

Giest and Klievink [3] studied the case of the State of Michigan where the Automated Fraud Detection System (MiDAS) was implemented, and the Dutch Childcare

Allowance case. In the US, the implementation of the AI system led to about a third of the Unemployment Agency staff being laid off. In the Dutch case, incremental changes affected the structure of the organisation, along with radical changes in the work of bureaucrats that had led to wrong decisions in the Dutch childcare subsidy.

Based on these observations, we see new organisational structures and human resource coordination in both of these two use cases, definable as administrative process innovation. In Michigan, we also see a conceptual overall breakthrough where there's a new way to approach the problem of fraud, through an AI system that can handle the process more accurately, faster, and more effectively than a human being. These observations speak of a more radical innovation in the latter case, and of only successive and incremental innovations in the Dutch case. This is in line with research that emphasises how the implementation of information technologies in public sector organisations is a complicated process and where change emerges incrementally [18], cited by [3].

In such scenarios, some envision new bureaucratic roles that correspond to digital processes. Bovens and Zouridis [19] suggest that there will be three groups of employees: (1) in the data processing process, such as system designers, lawyers, and system analysts; (2) managers in control of the production process; and (3) the "interfaces" between citizens and the information system, in the help desk, and also lawyers who deal with complaints. System designers, public policy implementers, and IT specialists are the new equivalents of bureaucrats [19]. However, some bureaucrats will be needed in the implementation of public policies and decision-making [20], and the ethical concerns involved. Thus, algorithms will just be additional decision support tools rather than becoming autonomous agents on their own, with their own ethics.

The Michigan Integrated Data Automated System (MiDAS) [3] has also been widely discussed not only in the media but also in legal proceedings, such as a class action lawsuit and Auditor General Reports, and new legislation has been passed to accommodate changes in the way Michigan Unemployment deals with MiDAS regarding claims of fraud.

3.1 Michigan Automated Integrated Data System (MiDAS)

In the USA, in the State of Michigan, data mining techniques were implemented for the automatic detection of program fraud in support related to food stamps and unemployment benefits, with the aim of reducing operating costs and targeting fraud in insurance claims, the so-called Michigan Integrated Data Automated System (MiDAS). This system in a fully automated way identified 48,000 fraud claims of unemployment benefit recipients—a five-fold increase over the previous system [21]. Some Michigan residents on food assistance, for example, were automatically disqualified by the system [22]. Automatic fraud determinations also occurred if recipients did not respond to the questionnaire within 10 days or if the MiDAS system automatically considered their responses unsatisfactory [22]. Claims from beneficiaries were systematically denied by the system, which failed to provide evidence to support MiDAS' accusations of fraud [23]. This system was presented by [3] to highlight its flaws on detecting fraud in the referred support claims.

Therefore, both cases are very similar regarding the objectives linked to the implementation of both AI systems, which is to streamline the process – making it more

efficient and cost-effective – and to identify fraud more reliably. But, in the case of Michigan, this objective goes a step further by pointing to very limited or even non-existent human intervention in the process, to the point of dismissing employees for specific tasks and replacing them with an automated decision-making system, resulting in major changes in the way bureaucratic work is defined [3].

3.2 Dutch Childcare Subsidy Fraud Evaluator Algorithm

The Dutch Childcare Allowance case was widely publicised in the media and extensively documented and reviewed by supervisory bodies: the Data Protection Authority, the Auditor General, and the Dutch Parliament. This system was aimed at classifying the risk of systematic fraud, based on an algorithm that was fed with examples of correct and incorrect applications in order to learn how to 'recognize' those that pose a risk of fraud. This innovative system sought efficiency, but also with the political objective of repressing fraud. There was a strong organisational incentive to rapidly detect fraud, which, counterproductively, provided a disincentive to look critically at system results. It was set up in a hurry and the procedures, its work instructions poorly adjusted, and with little specialisation at various levels, to cite civil servants who worked at the agency. The focus on efficiency and effectiveness in combating fraud was not accompanied by an adequate organisational structure [3, 24]. As a consequence, the Dutch tax authorities used algorithms that wrongly identified fraud in the attribution of support for access to day care centres. The alleged offenders, several with an immigration background, were summoned regarding the financial support, which caused great financial and psychological difficulties for the families involved. The data protection authority would come to conclude that the data processing used by the AI system was discriminatory [3].

3.3 Public Management Technologies and Models

New Public Management (NPM) began in the 1980s in OECD countries, due to the realisation that Public Administrations represented an excessive burden and a tendency towards inefficiency. This model seeks to reverse the trend towards an increase in the number of workers controlling public expenditure, resorting to private services to carry out public functions that do not collide with the social function of the State. It also promotes automation, via information technologies [25]. According to [4] this model of public management is closely associated with ideas of the managerial style of the private sector, a strong orientation towards customer service, and the use of less hierarchical organisational control mechanisms. This paradigm is characterised by three characteristics, namely: disaggregation by dividing the hierarchies of large public organisations; competition allowing for multiple forms of provision developed between suppliers, and incentive characterised by specific pecuniary performance incentives for personnel.

Ojo et al. [4] refer to three main successors for NPM: Digital Age Governance (DAG), Public Value Management (PVM), and New Public Governance (NPG), all with a common aspect – the centrality of digital technologies. DAG highlights the influence and impact of information and digital technology developments on public sector management, with emphasis on 1) transparency; 2) the use of platforms to transmit information; and 3) creation of service shared centres. Within the scope of transparency, open access

to information in a specific area or program of public interest is promoted (the so-called targeted transparency). The digital platform aims to increase citizen engagement by fostering civic participation, better engaging citizens to crowdsource ideas about policies and programs, plus opening more channels for citizens to voice their opinions about government services and programs. Shared services seek to achieve greater effectiveness and efficiency [4, 5].

PVM has as its paradigm the achievement of public value, attained by deliberation of citizens' representatives and government staff, surpassing the sum of individual preferences of public service users. It is based on a strategy aimed at creating public value. It seeks the use of real-time data to protect entities and services, favouring autonomy and entrepreneurship, and values innovation through creativity and out-of-the-box ideas. The PVM seeks through political management to ensure legitimacy in the assumption of value propositions. It resorts to social networks to stimulate opportunities of public value; it also seeks the realisation of politically empowered citizenship; plus seeks to learn across multiple levels and audiences, cultivating a broader vision of well-being for all [4, 26].

The NPG public management model is based on institutional and network theory [27]. NPG is achieved through networking and collaboration, public-private partnerships, and citizen involvement in public decision-making. Co-production/co-creation with citizens is fostered by collaboration in networks. In public-private partnerships (PPP), both sectors share risks and resources to achieve value. Through PPP innovation is also attained, given the different origins of the partners. A summary of the characteristics of paradigms subsequent to New Public Management (NPM) is found in [27].

Disruptive digital technologies such as AI are enacted in public management models associated with the era of digital government and public value management. We found no case where AI was deployed in the context of NPG. This may be due to the nature of AI solutions currently being deployed [4].

The European Law Institute (ELI) conceived the "Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration" [28], considering that PA is confronted with specific challenges resulting from the implementation of AI algorithmic decision-making systems and machine learning. Use of these techniques poses specific problems relating to the principle of good administration. Furthermore, issues such as transparency, accountability, compliance, and non-discrimination are particularly relevant in the context of PA. This European project led to the development of such ELI Model Rules [28]. Although inspired to some extent by EU legislation, it is compatible not only with existing EU legislation but also with draft legal provisions, the ELI Model Rules have been designed in such a way as not to depend on EU law. In this sense, they can serve as inspiration for national legislators (even outside the EU), for governments and PA. Algorithmic decision making can be approached in several ways. The central idea behind the Model Rules [28] is an impact assessment, adopting an approach that distinguishes between high-risk systems that warrant an impact assessment (in its Annex 1) and low-risk systems (in its Annex 2) where such an assessment is not warranted, and even systems that cannot be easily classified *ex ante* as belonging to Annex 1 or Annex 2, in which case a risk assessment will be made if screening reveals that the system poses at least a substantial risk [28].

3.4 Public Administration and Technology Adoption

A crucial role is reserved for the PA in the adoption of AI, considering its double condition as a user of the enormous potential of these technologies, but also with regulatory functions of AI, defining procedures, rules and competences, and particularly with regard to the Ethics of AI, establishing the minimum conditions for the private sector and citizens to deploy and use it in accordance with ethical requirements, but this is still lacking, hence our call to attention. EU Member States have given greater emphasis to the regulatory role of PA or, at best, to a facilitating role, that is, as the entity that defines the basic conditions for the ethical use of AI by private entities and citizens, but not by the PA itself. PA's role of "first buyer" [29] and direct beneficiary of AI is somewhat neglected. Most AI literature views the government as a regulator. Discussion of the role of PA from the point of view of an AI user is scarce, although PA is increasingly becoming a significant user of AI [30]. [31] refers it is essential that this issue be taken on with greater accuracy, considering the various areas in which PA operates, from health to education, and other areas with a wide range of services and entities that represent an essential market for AI, not only as buyers, but as well as disseminators of these technologies in the economic and social sectors. In this regard, [32] understands that PA will be the area where the opportunities and challenges of AI will constitute a priority on the agenda, with repercussions felt by those administered. In the opinion of [29], the key to success is to gain the trust of users in the processing of their data. To reach this opinion, overall, 686 AI use cases in the public sector were collected and analysed as reported in [33, 34].

The adoption of AI in PA procedures has the potential to bring greater efficiency and effectiveness in the provision of services to companies and citizens, increasing the level of satisfaction and confidence in the quality of the public service [4, 29]. The main applications of AI in this context include process automation, virtual agents and speech analytics, predictive analytics for decision making, sentiment analysis, and document review [4, 35]. Madan and Ashok [36] highlighted two specific technologies: machine learning (ML) and natural language processing (NLP), considering that these two technologies characterise most AI applications in PA, with reference to the cross-case analysis of [36] and to the AI case study archive of the European Commission and Joint Research Centre.

In this perspective, PA can also be seen in its distinctive aspects to the viewpoint of the private sector. For the pursuit of public interest differs from the profit motive that presides over private entities. In the study and implementation of PA AI systems they must not just aim at economic benefits, in accord with what was previously exposed concerning the pursuit of the public interest, and its human development perspective.

4 Ethics in Public Administration

The functional environment of public organisations is dynamic, full of unpredictable and ambiguous international events, changes in power, domestic policy challenges, and technological changes that affect the functioning of PA entities. Currently, workers in public functions increasingly work in a world characterised by volatility, uncertainty, complexity, and ambiguity, resulting in multifaceted and contradictory requests that can

configure dilemmatic situations [37]. In this scenario, values, integrity, conduct and ethics can motivate and support state agents.

Quotes [38]: "Savater [39] assumes that ethics is a kind of moral synthesis. An imagined ideal of collective existence, a social community with a united and coherent character." Ethics is thus a natural concern in the discussion about the real role of PA decision-makers in exercising the discretion of the powers entrusted to them, as well as in the pursuit of the public interest by holders of political office [40] and striving to ensure that algorithms respect and embody such ethics [41].

Interest in ethics in PA has been increasing with the introduction of the principles of NPM [25] in public sector bodies and services [1, 42, 43], because of the adoption of notions such as efficiency, obtained results, and economy, sparse in ethical content and conducive to pressures in the public service, causing doubts/confusion about the application of ethical standards. In this way, it is important to question why these ends, also implying a reflection on the means and relationships that are established within entities [44]. The importance of ethics is revealed, from the outset, by the vision and perception of the reason for being of PA and how its action should be undertaken in terms of demands and adaptation to change. The existence of a current crisis of values is often mentioned in different environments and contexts. Indeed, the aim of a civilization of lightness means all but living lightly, at present. Although the weight of social norms has lightened, life feels heavier. In unemployment, precariousness, instability in couples, and duration of overwork [45].

Lipovetsky [45] highlights the importance of ethical values and ideas, realising that despite the findings of extreme individualism, we are not located at the zero degree of values. The sense of moral indignation has by no means been eradicated, as our societies reiterate a stable core of shared values: human rights, honesty, respect for children and rejection of violence and cruelty. Although it refers to the inconsideration of sacrifice in the old-fashioned way, it does not mean the disappearance of the spirit of responsibility and solidarity.

It is in a context that considers ethics in public entities within a general scope that the ethical issues of AI can be more accurately framed. Notwithstanding that their specificity might be seen from an autonomous perspective, the ethical tools that already exist in PA, or that may come to be contrived, will facilitate the dissemination and application of this new burgeoning area of ethics [6].

5 The Ethics of the Machine

In the framework of ethical issues in AI systems for PA, it is important to summarise the ethics of the machine in general terms to outline possible adaptations to the public sector. The ethics of the machine, an expression coined by [46], has as its main objective to conceive a machine oriented with compatible ethical principles in the scope of the decisions it takes and on the possible consequences of its action. Fundamental human rights complemented with codes of ethics are considered capable of mitigating the risk of AI to evolve in an opposite direction to human values, by identifying risks, priorities, vulnerabilities [41, 47–51]. The Universal Declaration of Human Rights (UDHR), although not binding, is part of the legal system of democratic states.

For [52] transmitting ethics to robots by observing humans faces a prior complication—the human difficulty in determining which human attitudes collide with ethical principles, and how to resolve our contradictions. The greatest benefit may lie not in what we teach the machines, but in what we learn from teaching them. [41, 50, 51] also mention the difficulty of the lack of a universal ethics. On the other hand, our moral conduct is essentially instinctive and only in more complex situations it is necessary to think about right or wrong in terms of specific difficulties or discomfort.

The issues of weaponry in general which uses AI and the dangers arising therefrom with ever more powerful and more efficient weapons are dealt with in some detail by [53], who warn of the central paradox that the greater the digital capacity of a society, the more vulnerable it becomes. Computers, communications systems, financial markets, electricity networks (and the digital command and control systems they depend on)—even the functioning of the democratic system—involve systems that are, to varying degrees, vulnerable to manipulation or cyberattack. In a conflict situation, the most extreme form of protection may involve cutting the network connections, taking the systems offline. Disconnection could become the ultimate form of defence. Consequently, ethical principles need to be considered for addressing issues and situations that the law still has difficulty in resolving.

The existence of areas of absence of legality; the tendency towards the universality of ethical postulates; the anticipation of legislators' responses; the assistance given to the interpretation of legal rules; the capacity to integrate loopholes in the law; the increased efficiency of bonds arising from voluntary adoption—are some of the foundations that support permanent ethical intervention [54].

Bostrom and Yudkowsky [55] warn of the necessary care with applied ethics in contexts that are very different from those of the human condition. Common normative precepts are subject to conditioning in the face of different empirical conditions, and it is also important to consider whether these precepts are suitable for future hypothetical cases in which their preconditions might be considered invalid.

What is relevant is that an agent, when faced with a problem, can have the ability to generate hypotheses and choose the solution that seems most appropriate to him/her, through a causality supported by their experiential memory, by previous decisions, and by accumulated preferences [41].

Hagendorff [56] and Jobin et al. [57] selected documents with references to ethical principles and guidelines in the field of AI, in the public and private sectors, immediately indicating that the ethical challenges of AI are transversal to society. These authors listed the main ethical issues they identified in their research: transparency; justice and equity; non-maleficence; non-discrimination; responsibility; and privacy, are the dominant ones. Jobin et al. [57] emphasise that the principle of transparency prevails with approaches related to explainability and interpretability, with a main focus on the use of data, the human interaction with automated decisions, as well as data system applications. Transparency is presented as a method to minimise harm, but also as a factor that fosters trust, as well as a promoter of freedom and autonomy. The above references have been the object of criticism that considers transparency as a pro-ethical principle, a condition that can affect or harm other practices or ethical principles, namely set forth by [58].

5.1 Contributions of Ethical Perspectives in the Implementation of AI Systems

In an investigation at Instituto Superior Técnico, University of Lisbon, utilitarian ethics were examined, and general egalitarianism found questionable in an AI system that does not confer a main advantage to its owner. It was found that either AI offers an individual advantage to the owner, or it will not be adopted. The consequent lack of interest in its purchase will have negative effects on production, restricting research and development [59]. Hence there is an onus to incorporate in AI systems ethical values proximate to their owners, though without disregarding general social values.

In this respect, the ideas of Rawls [60] have been defended for the field of AI, for example by [61]: "A Theory of Justice could help explore known problems of data bias, injustice, liability and privacy, in relation to machine learning and AI applications in government." [62], in his approach to the ethics of AI, considers that in Rawls's conception of justice, the veil of ignorance can be replaced by a much more natural condition of prudent selfishness in a finite world. The American contractualist Rawls is also cited by [63] in reference to a new social contract with the intervention of algorithms. Leben [64] describes a Rawlsian algorithm as an alternative to a utilitarian solution, which will assess the probability of survival for each person in an autonomous vehicle, for each action, calculating which action would obtain agreement starting from an original position with fair negotiation. Rawls [60] is also mentioned, for example, at the European Union level regarding the need to use AI in decision-making which can overcome human judgement. It is questionable whether a human will still contribute to a stronger protection of data subjects, or whether the better performance of machines even regarding the political and legal values at stake; for example, guaranteeing fair equality of opportunities, as defended by [60], thereby making human intervention redundant or dysfunctional [65].

Rule-based ethical theories, such as the Kantian one [66], are considered promising for machine ethics due to the structure of their judgments. According to a formalist interpretation of Kant's categorical imperative—"Act as if the maxim of your action should become, by your will, a universal law of nature" [66] —a machine might, for example, place prospective actions on top of traditional maxims (forbidden, permitted, obligatory), in the top-down approaches to the development of programming agents in which Kant's utilitarian ethics and deontological ethics are likely to be applied [65].

The principle of responsibility theorised by Hans Jonas was at the origin of the construction of a new understanding regarding the ethical aspects resulting from the position technology assumes in society. He considered responsibility as a basic ethical principle of human action, given the issues raised by technology, defending how the theory of responsibility can be seen as an ethics par excellence [67]. This author refers to an imperative for the new technological types of human action, that addresses the new type of object of action. Jonas [67] states: "Act in such a way that the effects of your action are compatible with the permanence of an authentically human life in society on Earth." The new imperative is addressed to public policy, as opposed to Kant's categorical imperative being addressed to the individual subject [67]. Alexandre [6] found that such Jonas's principle of responsibility is still not accepted within the scope of machine ethics. However, it seems to us this gap may be filled in the near future, as a result of his intergenerational approach, which is concerned with the legacy for the next generations.

6 European Union: An "Engine" of AI Ethics

The European Union (EU) has been looking for foundations for the implementation of an ethical AI, publishing several documents that justify an analysis not only for their pioneering spirit, but also for the repercussions that it may have in European countries. The EU's concerns with ethical aspects and fundamental human rights are evident, namely in the promotion of a human-centred AI defending human values as an "engine" of development and economic and social progress, seeking in particular to mitigate existing inequalities.

Alexandre and Pereira [68] have addressed the main documents produced by the European Union highlighting ethical concerns with the implementation of AI in PA. Of the documents analysed by these authors, we highlight three:

– In 2019, the High-Level Expert Group on Artificial Intelligence (AI HLEG), which was appointed by the European Commission [69], released the Ethical Guidelines for a Trusted AI, the first institutional document emphasising at planetary scale the importance of ethics in the development and implementation of AI. In this document, PA was also considered a beneficiary of improvements that AI systems present to the efficiency of public services, in their provision of public goods and services to society [68].

– In the diagnosis made by the "White Paper on AI—An European approach towards excellence and trust," published 19 February 2020, on the use of AI in the public sector, where the innovation capabilities and greater efficiency of services are also highlighted, with express reference to hospitals, transport services, public utility in general, financial supervisors, and other areas of public interest [7, 68].

– The Proposal for a Regulation, of the European Parliament and of the Council, from 21 April 2021, which establishes harmonised rules on Artificial Intelligence, known as the Artificial Intelligence Act (AIA) [69], constitutes the first institutional proposal that aims to regulate AI, openly challenging those who believe that the law should not regulate emerging technology, a very common stance in Silicon Valley [70]. This proposal for a community regulation highlights a change in the Commission's narrative: whereas in the White Paper one could identify the reversal of EU priorities, previously assuming a global competition neglecting fundamental rights, the "Proposal for a Regulation" advocates to ban AI practices which, via high-risk AI systems, may violate the European Union values and fundamental rights enshrined in the Charter of Fundamental Rights of the European Union, while proposing softer provisions for lower- and medium-risk AI systems. This proposed regulation foresees the future rules regulating AI to be supervised and monitored by national authorities. This may be an opportunity for member states not yet having created an entity with competences in the field of ethics in AI to create, as proposed in [41], a "National Ethics Commission for AI" to oversee the foreseen regulatory structures by Brussels, with a higher-up competence in the field of ethics [68].

6.1 European Strategies

According to [29, 34] the AI Strategies of EU Member States regarding the public sector show ethical concerns, with some expressing the intention to develop an ethical framework to guide the implementation of AI in the public sector, seen as a contribution to

establishing trust between workers in public functions and citizens, ensuring the quality and compliance with ethical values of the AI used by the PA. At the level of legal regimes, legal reforms are referred to with the aim of streamlining the development and use of AI. Some countries have noted difficulties in exchanging data between public entities and the private sector, and thus intend to create "regulatory structures" to facilitate this exchange. A significant number of these reforms will be specific to each sector, with emphasis on health [29].

In the European Union, in February 2022, national AI strategies had already been published in 23 Member States, also including Norway, establishing conditions for the development and acceptance of AI. The public sector's pioneering spirit is affirmed in the 2021 Review of the Coordinated Plan on Artificial Intelligence. It affirms the potential of AI in modernising the public sector itself, with the capacity to: "(i) automate simple and repetitive cognitive activities freeing-up labour time for more high-value activities; (ii) increase the predictive capabilities, enhancing data-driven decision-making; and (iii) support user-centric service personalization, increasing the effectiveness of public service delivery" [33].

7 Obstacles in Public Administration to the Implementation of AI

7.1 Organisational Culture

At the level of public services and entities, an organisational culture is decisive in the implementation of AI and its acceptance. In organisational dynamics, it is important to combat the inertia inherent in a routine rigidity that inhibits change and the development of new capabilities, but also the inertia that results from the constant change caused by electoral cycles, in democratic regimes that may lead to successive political changes, where concerns about re-election in the latter part of a term usually condition the decisions of political office holders. The influences of other government departments cannot be overlooked, as well as pressure from civil society and the media. As a result, PA, in addition to its internal difficulties, is also part of a turbulent and volatile external environment [36]. Inertia is considered a critical factor, due to the scarcity of resources for pilot and/or innovative projects, but also due to the difficulty in finding AI specialists. Inertia may also result from the rigidity of bureaucratic factors, centralised decisions, poorly trained workers, and too from attitudes contrary to the sharing of data within public entities and with outside departments [36].

At the PA's internal level, facilitating factors for the implementation of AI are an organisational culture with innovative and dynamic characteristics, favouring experimentation to face the risks of these technologies. Transformational leaders will motivate their workers to change and will seek to influence design as well as interaction with other agencies, services, and departments [30, 71, 72].

The Technology-Organization-Environment (TOE) structure of [73], as mentioned in [36], is indicated to explore the adoption of technology in different environments, due to the identical importance that it gives to both organisational and technological contexts in the implementation of technology in organisations.

The processes of diffusion and adoption of technological innovation according to [74], operate for long periods of time, can be understood in multiple analyses, and at

different levels of aggregation. They highlight the inseparability of human values, in the complex interactions between technological innovation, people, scientific concepts, aspirations, and consequences.

7.2 Contribution of Public Procurement

According to [6] the clear definition of AI ethics issues should be placed *ex ante*, in the acquisition and development of AI systems by the PA, with a clear specification in the terms of reference. The public entity may resort to preliminary market consultation, listening to potential suppliers as to the conditions they offer to guarantee ethical requirements. Thus, in the formulation of public policies with a view to implementing these technologies, the Government's purchasing power can play a key role in this matter, as referred in [7] and [29] with a determination of ethical criteria in the acquisition requirements that ensure private companies designing AI systems adequately meet public standards. This requirement can also contribute to the dissemination of ethical AI in the private sector.

Desouza et al. [75] propose in this context an agile acquisition process that allows for iterative development lifecycles through the acquisition of hardware and software in stages, ensuring that early access to knowledge of a sector focuses on problem definition, rather than developing detailed solution specifications.

7.3 Data – What Balance Between Security, Privacy, and Innovation?

The PA, in the performance of its activities, collects a huge amount of data from citizens, superior to most private entities [4], and data collected, for example, in the areas of taxation, health, education, and social security are particularly sensitive. Data privacy and security have become increasingly important, resulting in the publication of legislation of which the EU's General Data Protection Regulation (GDPR) is the highest exponent [76].

Madan and Ashok [36] refer that the accessibility of data and its use by governments for purposes other than those that were collected raise serious concerns related to privacy. On the one hand, the use of data can lead to superior public policy and service delivery about duty-oriented and service-oriented public values. However, at the same time, it undermines the social public value of privacy. [77, 78] state that these data provide information that may allow classifying citizens into clusters of micro populations. Thus, the accessibility of data and its use by governments for purposes other than those for which it was collected raise serious privacy concerns. There is a balance that needs to be achieved between the protection of personal, sensitive, and confidential data and, on the other hand, the potential of data collected by the public sector for the design and development of AI systems aimed at the public services, which manifest aptitude for the improvement of public services that will reflect on the citizens' own living conditions, thus contributing to progress.

A particularly sensitive issue is predictive policing, that is to police individual behaviour in order to attempt to predict rule infringement by the individual. The main ethical issues gravitating around data selection and machine biases, such as "visualisation and interpretation of forecasts, transparency and accountability, time and effectiveness

as well as the problem of stigmatisation of individuals, environments and community areas" [79].

7.3.1 Privacy and Data Collection

The concept of privacy has evolved since the dawn of the fourth industrial revolution, in the permissions that users grant to electronic platforms and in the exposure that users make of their lives on social networks. Even so, privacy is an aspect that inspires concern in the implementation and use of AI in public services. Ensuring respect for citizens' privacy is considered one of the essential aspects for winning trust in AI and for it to be successful [6].

In the interviews carried out by [6] a specialist's reference was made to the balance that must be achieved between data privacy and the need to compromise in terms of solutions that are adequate to guarantee public health and, ultimately, save human lives. It was mentioned that in a pandemic situation, like the recent pandemic situation due to the spread of the Covid 19 virus, several rights can be limited. The right to privacy is almost considered evangelical, seen as if it were a greater right than the right to life itself. Therefore, it is necessary to establish a discussion regarding the limitation of the right to privacy in a pandemic context. The specialist cited the example of South Korea, which had a lot of cases in the first SARS. To prevent the spread of Covid 19, it used data such as economic transactions with credit cards, mobile phone data, and Bluetooth communication so as to test people and quarantine them before transmitting the virus. This decrease in privacy has occurred despite South Korea having a democratic regime. For this interviewee, the issue of privacy was identified as the most important in the PA, and he also referred that although research on data from people with cancer is regulated, preventing the identification of the patient, a detailed analysis of the data allows this identification.

8 Discrimination and Prejudice

Prejudice, discrimination and racism, or biases related to gender equality, are risks involved in any economic or social activity. Human decision-making is not immune to error and bias. However, the same prejudice in the field of AI and particularly in PA can have a much greater effect [6]. Within the scope of AI, discrimination is identified "on grounds of sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation" [7]. Berryhill et al. [80] refer that in the very conception of AI tools, the lack of diversity is reflected in the technology industry. For example, only about 19% of research related to AI is authored by women; on the other hand, the number of AI publications authored by women has stagnated since the 1990s.

The models of recidivism risk introduced in the US judicial system, designed by data-fed algorithms, demonstrate racial discrimination against African-Americans [81–83]. They make it possible to assess the danger of recidivism that each convict represents, with the conviction that this assessment is more accurate than a judge's off the cuff guess [84]. The LSI–R questionnaire focuses on ten different topics, addressing questions about the inmate's birth, education, family, neighbours, and friends, classified with different

indices depending on their importance, as mentioned by [85], in the publication "LSIR – What is the Level of Service Inventory-Revised Risk Assessment?", edited by LLC, Trial Lawyers. For [82], the questionnaire does not directly identify the race, as this would be illegal, but with the quality of the details that each prisoner provides, this illegal question becomes almost superfluous [86].

An independent investigation carried out by ProPublica (a non-profit entity doing investigative journalism), analysed COMPAS—Correctional Offender Management Profiling Alternative Sanctions—concluded that defendants of African descent were much more likely to be incorrectly assessed as having a higher risk of recidivism compared to Caucasian individuals, as for the latter was found an incorrect tendency towards low-risk flagging.[1]

In the interviews carried out by [6] of specialists in the world of AI, from university professors, civil society agents, to public administration leaders, the following were highlighted:

Ethical issues related to discrimination that AI systems are still unable to avoid, resulting from the bias of algorithms that produce discriminatory decisions, were the main issues identified in this context.
Neural networks were mentioned due to the lack of language concepts capable of explaining the decisions they make.
The biases result from the actions of the humans that are at the origin of the data which will contribute to discriminatory actions. If the methods used cannot avoid this bias, then the mistakes of the past will be repeated and reinforced.
The need for communication of the possible biases, by those responsible for the design, to the end adopters of the systems.

In machine learning, the under- or over-representation of data for some groups may occur, so the careless use of data in AI training can cause or perpetuate discrimination and inequalities that already exist in society. AI itself can contribute to sexism and gender stereotypes through tools that enhance such behaviours, e.g., most virtual assistants are given female names and personalities associated with sexist/stereotypical "female reactions" and some hardier physical robots (e.g., rescue robots) were given male forms [87].

Discrimination within the scope of AI must also be seen in terms of unequal access, considering citizens with insufficient resources and/or knowledge to establish a technological connection with Public Administration services and bodies, as they are in a situation of digital exclusion. Electronic administration cannot serve as a factor of discrimination between citizens with digital capabilities and digital illiterates. The obligation, in the most complicated periods of the pandemic, to hold basic, secondary and higher education classes through the Zoom and Teams platforms, faced with the difficulty of connecting to the internet due to the lack of necessary equipment, either by students or by some teachers [88, 89]. In an interview conducted by [6] it was mentioned that, of

[1] Interested readers can consult the work and methodology of this entity at: https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis.

course, knowing how to send an email or use WhatsApp is not enough for a full-fledged insertion in the world of technology.

Bias in AI is verified when machine learning can lead to discrimination against people, specific groups, generally marginalised by gender, social class, sexual orientation, race, or religion. It may originate from prejudiced attitudes in the very design of the AI system, and/or be due to the use of non-representative, inaccurate or simply wrong training data. These are cases of discrimination, legally defined as the unfair or unequal treatment of an individual (or group) based on certain protected characteristics (also known as protected attributes) such as income, education, gender, or ethnicity [24, 90].

Service providers, both public and private, must consciously address issues of prejudice and discrimination, ensuring that behaviours, experiences, and views that represent the diversity of the population have been considered, cf. Committee on Standards in Public Life [91].

9 Explainability/Transparency, Transparency and Responsibility

In terms of explainability, the opacity of the algorithms that constitute autonomous AI systems that make decisions, but having difficulty in explaining the respective grounds, is a major concern. Users expect to understand or else be explained the grounds for those decisions. In the absence of these grounds, recipients are limited in any claim they wish to present. [92] refer that at present machines are not "good storytellers", meaning that they cannot explain the gist and flow of their reasoning and conclusions in a language readily understandable by human users.

According to [8], opacity can be: (i) intentional to protect intellectual property [93]; (ii) illiterate, where a system is only understandable to those with technical skills; and (iii) intrinsic, the complexity of the system makes the understanding of its decisions difficult for any human being.

Olsen et al. [94] conceive that the implementation of AI in PA raises some concerns that may cloud legal measures, delaying the implementation of these systems. The first concern is the loss of control over systems and therefore a clear link to accountability when decisions are made. In the exercise of discretionary power, the agent assumes responsibility for the decision. The fear is present even when the AI system is used in conjunction with the human or supervised by the human, yet the deference to the machine creates a vague sense of responsibility for the decision. The second fear is the loss of human dignity, with the reduction of humans to mere "cogs in the machine". Taking away the ability to understand and communicate freely with another human being can easily lead to alienation and loss of human dignity. Finally, the difficulty of using data that will lead to false and discriminatory decisions was also identified.

A hybrid system was proposed by [94]: the algorithm using machine learning in PA would be used to produce drafts of decisions that can increase efficiency in services or organisations without reducing quality, provided that the data submitted to machine learning presents a sufficiently high volume and learning has a point of reference in well-founded decisions. The authors state that trainees also learn from previous decisions and that humans differ from algorithms because these tend to be more rigorous than humans.

Actions can be filed in the courts against procedures with automated decision-making, due to the opacity of the systems preventing the understanding of the reasons

for the decision [8]. These systems, in particular when opacity is intrinsic, need to be adapted to the need to justify acts in force in PA, so that their implementation in the public sector does not result in the annulment of acts whose foundations are unintelligible.

Regan et al. [78] conceive that the opacity of AI systems in whatever PA service challenges the traditional responsibilities of administrators, regardless of substantive policy, or whether decisions directly involve citizens. Therefore, when considering the use of AI systems, the manager should take into account:

The likely effects of an AI system on its oversight, as well as on unintelligible material in its decision making.
Determine the regime of responsibility resulting from the decisions of such systems.
Determine acceptance levels of AI autonomy and the extent to which they should be changed.

Transparency and explainability in AI-based decisions can elicit greater trust from both PA officials and citizens. However, the downside of increased transparency is the ability to manipulate the system for private reasons [4, 95].

The implementation of AI and respective ethical guidelines is a multidisciplinary process requiring contributions, among others, from technology, ethics, statistics, law, social sciences, legislators, journalists, including politicians and the population at large [43, 80]. Depending on the respective application, contributions from sociologists, psychologists, doctors, or others with experience in the respective area will also be necessary [96].

10 Conclusions

Public Administration (PA) services and entities as advanced technology users, and namely AI, represent a particularly important position in the market, given the number of areas that comprise it and the organisations that gravitate in its sphere, thus being considered as the main promoter and buyer of technological systems. The PA is also assigned the role of regulator, responsible for protecting users in terms of security, but also for defining the basic conditions, in particular for the growing preoccupations with ethical use of AI systems by private entities and citizens.

The public interest that characterises the public administrative function, as the main differentiating factor from private activity, must justify the search for technological solutions that contribute to improving the human development index (HDI), allowing for longer, healthier, and more creative lives, in line with the human-centricity aims of Industry 5.0 and Society 5.0.

Technology, namely AI, is producing changes in technical areas, in communication, and in decision-making. Public services have been endowed with greater effectiveness, efficiency, and transparency.

However, the implementation of AI in PA as a user has received less attention from political power, compared to the regulatory role of AI in terms of security and the conditions for the implementation of AI ethics.

Some models for evaluating the risk of fraud in the attribution of social benefits, such as the subsidy for day-care centres in the Netherlands, or the unemployment fund in the

State of Michigan, in the United States, were poorly implemented, with flaws in the assessments carried out, and in the inability to present adequate grounds in identifying fraud risks.

The interest in ethics in PA has been increasing with the application of New Public Management principles such as efficiency, results, and economy, which are however sparse in ethical content and conducive to pressures on the service. The importance of ethics is revealed by the vision and perception of the PA's reason for existing and how its actions should be enacted in terms of diligent demand and adaptation to change.

It is in a context that considers ethics in public entities in a more general context that the ethical issues of AI can be framed with greater accuracy, albeit their specificity may be envisaged from an autonomous perspective. The AI ethical tools that already exist in the PA, or that come to be conceived in it, will facilitate the dissemination and application of this new area of ethics.

The ethics of the machine is faced *ex ante* with the difficulty arising from the non-existence of a universally accepted ethics, therefore it being necessary to define which human attitudes collide with ethical principles, and how to resolve our own contradictions. The ethics of AI has also been designed to fill in an absence of legal rules, aiming to anticipate responses from the legislator, but also contributing to the interpretation of legal rules, the integration of loopholes in the law, plus the increased efficiency provided by voluntary adherence bindings. Some AI systems have been designed respecting principles of ethical currents such as utilitarianism, contractualism, deontology, or virtue ethics. The new imperative defended in [67]: "Act in such a way that the effects of your action are compatible with the permanence of an authentically human life on Earth," addresses public policies, contrary to Kant's categorical imperative directed only to the individual, yet still forgotten by this new branch of ethics can positively contribute to the realisation of a more universal machine ethics.

References

1. Bilhim, J.: Teoria Organizacional – Estruturas e Pessoas, 7th edn. Instituto Superior de Ciências Sociais e Políticas, Lisbon (2013)
2. Shafritz, Jay, M., Russell, E.W., Borick, Chrisopher, P., Hyde Albert, C.: Introducing Public Administration. Ninth Edition, Routledge (2017) https://himia.umj.ac.id/wp-content/uploads/2021/08/Introducing-Public-Administration-by-Jay-M.-Shafritz-et-al.-z-lib.org_.pdf
3. Sarah, N.G., Bram, K.: More than a digital system: how AI is changing the role of bureaucrats in different organizational contexts, Public Management Review. https://www.tandfonline.com/doi/full/10.1080/14719037.2022.2095001 (2022)
4. Ojo, A., Mellouli, S., Ahmadi Zeleti, F: A Realist Perspective on AI-era Public Management – An Analysis of Mechanisms, Outcomes and Challenges of AI Solutions in the Public Sector. In: 20th Annual International Conference on Digital Government Research, Association for Computing Machinery. https://dl.acm.org/doi/abs/10.1145/3325112.3325261 (2019)
5. Tan, E., Crompvoets, J.: A new era of digital governance. KU Leuven Public Governance Institute. Wageningen Academic Publishers (2022). https://www.wageningenacademic.com/doi/pdf/10.3920/978-90-8686-930-5_1
6. Alexandre, A. C.: Novos Desafios da Administração Pública: Inteligência Artificial e Ética nos Sistemas Inteligentes com Autonomia. Masters dissertation, Instituto Superior de Ciências Sociais e Políticas, Lisbon (2022) https://www.repository.utl.pt/handle/10400.5/25585

7. European Commission: White Paper on Artificial Intelligence: A European approach to excellence and trust https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (2020)
8. Cobbe, J.: Administrative law and the machines of government: judicial review of automated public-sector decision-making. Legal Studies, vol. 39, pp. 636–655, Cambridge University Press, Cambridge MA, USA. (2019) https://scholar.google.com/citations?view_op=view_citation&hl=en&user=TsHYXWoAAAAJ&alert_preview_top_rm=2&citation_for_view=TsHYXWoAAAAJ:9yKSN-GCB0IC
9. Nwanisob, Beluchi C., Christopher, I. C.: Definition of Public Administration. Am. Int. J. Bus. Management (AIJBM) **3**(9), (2020). https://www.aijbm.com/wp-content/uploads/2020/09/I395661.pdf (2020)
10. Lippmann, W.: Public Opinion. Transaction Publishers, New Brunswick and London, Second printing (1998)
11. Haeberlin, M., Comim, F.: Todos e cada um de nós: o interesse público como critério de desenvolvimento humano. Revista Brasileira de Políticas Públicas, vol. 10. (2020)
12. Coulanges, F.: La Cité Antique. Librairie Hachette, Paris (1900)
13. Sen, A.: Development as freedom. Oxford University Press. (1999)
14. United Nations Development Programme: Human Development Report 1990: Concept and Measurement of Human Development. UN, New York. https://hdr.undp.org/content/human-development-report-1990 (1990)
15. United Nations Development Programme: Human Development Report 2010: The Real Wealth of Nations: Pathways to Human Development. Palgrave Macmillan, New York. https://hdr.undp.org/system/files/documents/human-development-report-2010-complete-english.human-development-report-2010-complete-english (2010)
16. Petropoulos, G., Marcus, J. S., Moes, N., Bergamini, E.: Digitalisation and European Welfare States. Bruegel, Brussels. https://www.bruegel.org/sites/default/files/wp_attachments/Bruegel_Blueprint_30_ONLINE.pdf (2019)
17. Erdurmazli, E.: Effects of Information Technologies on Organizational Culture: A Discussion Based on the Key Role of Organizational Structure. In: Süleyman Davut Göker (Ed.) A Closer Look at Organizational Culture in Action. https://www.intechopen.com/chapters/72534 (2020)
18. Meijer, A., Bekkers, V.: A Metatheory of E-Government: Creating Some Order in a Fragmented Research Field. Government Information Quarterly. https://www.sciencedirect.com/science/article/abs/pii/S0740624X15000568 (2015)
19. Bovens, M., Zouridis, S.: From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. Public Administration Review 62. https://onlinelibrary.wiley.com/doi/10.1111/0033-3352.00168 (2002)
20. Keiser, L.: Understanding Street-Level Bureaucrats' Decision Making: Determining Eligibility in the Social Security Disability Program. Public Administration Review 70(2):247 – 257. DOI:https://doi.org/10.1111/j.1540-6210.2010.02131.x (2010)
21. Gilman, M.: States Increasingly Turn to Machine Learning and Algorithms to Detect Fraud, U.S. News & World Report, Feb. 14. (2020) https://www.usnews.com/news/best-states/articles/2020-02-14/ai-algorithms-intended-to-detect-welfare-fraud-often-punish-the-poor-instead
22. Richardson, R., Schultz, J.M., Crawford, K.: Dirty Data, Bad Predictions: How Civil Rights Violations Im pact Police Data, Predictive Policing Systems, and Justice, New York University Review Online, Vol. 94:15 (2019)
23. De la Garza, A.: States' Automated Systems are Trapping Citizens in Bureaucratic Nightmares with their Lives on the Line. Time, May 28 (2020) https://time.com/5840609/algorithm-unemployment/

24. European Union Agency for Fundamental Rights: Bias in Algorithms –Artificial Intelligence and Discrimination. (2022), https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf
25. Carvalho, E.R.: Reengenharia de Processos na Administração Pública. Universidade de Lisboa, Instituto Superior de Ciências Sociais e Políticas (2001)
26. Sedlačko, M.: The five tasks of public value management: public value as a programme of administrative and societal democratisation. J. Public Adm. Policy **XIII**(2), Winter 2020/2021. (2020) https://sciendo.com/article/10.2478/nispa-2020-0019
27. Patapas, A., Raipa, A., Smalskys, A.: New public governance: the tracks of changes. Int. J. Bus. Social Res. **4**(5), (2014) https://econpapers.repec.org/article/lrclarijb/v_3a4_3ay_3a2014_3ai_3a5_3ap_3a25-32.htm
28. European Law Institute: Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration. (2022), https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSs_Used_by_Public_Administration.pdf
29. Misuraca, G., van Noordt, C.: Overview of the use and impact of AI in public services in the EU, European Union (2020) https://publications.jrc.ec.europa.eu/repository/handle/JRC120399
30. Kuziemski, M., Misuraca, G.: AI governance in the public sector: three tales from the frontiers of automated decision-making in democratic settings. Telecomm Policy **44**(6) (2020) https://pubmed.ncbi.nlm.nih.gov/32313360/
31. Alexandre, A., Costa: Porquê a Ética na Inteligência Artificial? Jornal i, 14/10/2022. (2022) https://ionline.sapo.pt/artigo/783201/porqu-a-etica-na-intelig-ncia-artificial-?seccao=Opiniao_i
32. Aydın, A.: Transformation of Public Administration in the Impact of Artificial Intelligence Systems (2020), https://medium.com/@sibervatandas/transformation-of-public-administration-in-the-impact-of-artificial-intelligence-systems-d564bfc40764
33. Ricart, J., Van Roy, R., Rossetti, F.V., Tangi, L: AI Watch - National strategies on Artificial Intelligence: A European perspective. EUR 31083 EN, Publications Office of the European Union, Luxembourg (2022). https://ai-watch.ec.europa.eu/publications/ai-watch-national-strategies-artificial-intelligence-european-perspective-2022-edition_en
34. Tangi, L., van Noordt, C., Combetto, M., Gattwinkel, D., Pingnatelli, F.: AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector. EUR 31088 EN, Publications Office of the European Union, Luxembourg. (2022) https://publications.jrc.ec.europa.eu/repository/handle/JRC129301
35. Wirtz, B.W., Weyerer, J.C., Geyer, C.: Artificial intelligence and the public sector—applications and challenges. Int. J. Public Adm. **42**(7), 596–615 (2018)
36. Madan, R., Ashok, M.: AI adoption and diffusion in public administration: a systematic literature review and future research agenda. Govern. Inform. Quart. **40**(1), 101774 January (2023) https://www.sciencedirect.com/science/article/pii/S0740624X22001101
37. Wal, Z.: Trends and drivers of public administration in the twenty-first century, Handbook on Corruption, Ethics and Integrity in Public Administration -, edited Graycar, Adam, Edward Elgar Publishing Limited, (2020) https://book4you.org/book/5679954/f5fc1b
38. Maia, A.: Ética e Integridade na Governação Pública – A corrupção e a fraude, e sua prevenção. In: Maia, A. (Ed.) Ética e Integridade na Vida Pública, Almedina. Lisbon. (2020)
39. Savater, F.: Convite à Ética. Fim de Século, Lisbon (2008)
40. Amundsen, I.: Introduction to Public Sector Ethics: In: Public Sector Ethics, Amundsen, Inge & Andrade, Vicente P. (Eds.). Universidade Católica de Angola, ResearchGate. (2009), https://www.researchgate.net/publication/275968636_Introduction_to_Public_Sector_Ethics_Introducao_a_Etica_no_Sector_Publico

41. Pereira, L.M., Lopes, A.B.: Machine Ethics: From Machine Morals to the Machinery of Morality. SAPERE vol. 53, Springer Nature AG, Cham, Switzerland. (2020)
42. Mozzicafreddo, J.: O papel dos Cidadãos na Administração Pública. In: Reforma do Estado e da Administração Pública. Gestionária, Instituto Superior de Ciências Sociais e Humanas, Lisbon. (2000)
43. OECD: Hello, World: Artificial Intelligence and its Use – the Public Sector. OECD Working Papers on Public Governance (2019) https://doi.org/10.1787/19934351
44. Soares, L.: M P: A Ética na Administração Pública. Instituto Superior de Ciências Sociais e Políticas, Lisbon (2014)
45. Lipovetsk, G.: Da leveza para uma civilização do ligeiro. Edições 70, Lisbon (2017)
46. Anderson, M., Anderson, S.: Introduction: Machine Ethics, IEEE Intelligent Systems (July/August), IEEE-org. (2006), https://ieeexplore.ieee.org/abstract/document/1667946?casa_token=NTsFmOCQ5xgAAAAA:a8eMjkNh6GM8e6Xz3UCJ0OSo5CXuhJtIb0IlG3MAVVs5zzAKRUp8XY4k1fflW6Z5y5KKqWG1ig
47. Organization Economic Cooperation Development: Recommendation of the Council on OECD Legal Instruments on AI (2019), https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449
48. Organization Economic Cooperation Development: Artificial Intelligence in Society, OECD (2019) https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm
49. Tavani, H.: Ethics & Technology – Controversies, Questions and Strategies for Ethical Computing. 4th edition, John Wiley (2013), https://futuresinitiative.org/futureofwork/wp-content/uploads/sites/216/2018/07/Ethics_and_-Technology-Controversies-Questions-and-Strategies.pdf
50. Nankervis, A., Connell, J., Montague, A., Burgess, J. (eds.): Springer, Singapore (2021). https://doi.org/10.1007/978-981-16-1614-3
51. Schawb, K.: Shaping the Fourth Industrial Revolution. Cologny, Geneva (2018)
52. Domingos, P.: The Master Algorithm, Penguin Books (2017)
53. Kissinger, H.: Schmidt, Eric & Daniel, Huttenlocker: The Age of AI And Our Human Future. Little Brown, London (2021)
54. Antunes, H.S.: Direito e Inteligência Artificial. Universidade Católica Editora, Lisbon (2020)
55. Bostrom, N., Yudkowsky, E.: The Ethics of Artificial Intelligence, Cambridge Handbook of Artificial Intelligence, edited by Keith Frankish and William Ramsey, Cambridge University Press (2011) https://doi.org/10.1017/CBO9781139046855
56. Hagendorff, T.: The ethics of ai ethics: an evaluation of guidelines. Mind. Mach. **30**(1), 99–120 (2020). https://doi.org/10.1007/s11023-020-09517-8
57. Jobin, A., Ienca, M., Vayena, E.: Artificial Intelligence: the global landscape of ethics guidelines. Health Ethics & Policy Lab, ETH Zurich (2019), https://export.arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf
58. Turilli, M., Floridi, L.: The ethics of information transparency. Ethics Inf. Technol.**11**, 105–112 (2009) https://link.springer.com/article/10.1007/s10676-009-9187-9
59. Fernandes, P., Santos, F., Lopes, M.: Norms for Beneficial A.I.: A Computational Analysis of the Societal Value Alignment Problem. INESC-ID, Instituto Superior Técnico, Universidade de Lisboa. https://doi.org/10.3233/AIC-201502 (2019)
60. Rawls, J.: A Theory of Justice. Revised edition, Harvard University Press and Google Books. (1971)
61. Grace, J., Bamford, R.: AI Theory of Justice: Using Rawlsian Approaches to Better Legislate on Machine Learning, Government, School of Advanced Study – University of London (2020), https://journals.sas.ac.uk/amicus/article/view/5161
62. Thomsen, K.: Ethics for Artificial Intelligence, Ethics for All. J. Behav. Robot. 359–363 (2019). https://www.degruyter.com/document/doi/10.1515/pjbr-2019-0029/html

63. Rahwan, I.: Society-in-the-Loop: Programming the Algorithmic Social Contract. Ethics and Information Technology 20 (2018), https://link.springer.com/article/10.1007/s10676-017-9430-8
64. Leben, D.: A Rawlsian algorithm for autonomous vehicles. Ethics Inf. Technol. **19**(2), 107–115 (2017). https://doi.org/10.1007/s10676-017-9419-3
65. European Parliament: The ethics of artificial intelligence: Issues and initiatives, European Union. (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf
66. Kant, I.: Fundamentação da Metafísica dos Costumes, Edições 70, Lisbon. (2007)
67. Jonas, H.: El Princípio de Responsabilidad – Ensayo de una Ética para la Civilización, Herder. (1995)
68. Alexandre, A.C., Pereira, L.: European Union Fosters Ethical AI in the Public Administration. In: de la Iglesia, D.H., de Paz Santana, J.F., López Rivero, A.J. (Eds.) New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence. DiTTEt Advances in Intelligent Systems and Computing, vol. 1430, Springer. (2022). https://doi.org/10.1007/978-3-031-14859-0_9
69. European Commission: Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative acts (2021). https://eur-lex.europa.eu/legal content/EN/TXT/?uri=CELEX:52021PC0206
70. MacCarthy, M., Propp, K.: LAWFARE, Home Hard National Security Choices, Law Fare (2021). https://www.lawfareblog.com/machines-learn-brussels-writes-rules-eus-new-ai-regulation
71. van Noordt, C., Misuraca, G.: Evaluating the impact of artificial intelligence technologies in public services: Towards an assessment framework.In: ACM International Conference Proceedings Series (2020), https://doi.org/10.1145/3428502.3428504
72. van Noordt, C., Misuraca, G.: Exploratory insights on artificial intelligence for government in Europe. Social Sci. Comput. Rev. 40(2), 426–444, (2020) https://journals.sagepub.com/doi/abs/10.1177/0894439320980449
73. Tornatzky, L.G.,Fleischer, M.: The Processes of Technological Innovation. Lexington Books (1990). https://archive.org/details/processesoftechn0000torn/mode/2up
74. Eveland, J., Tornatzky, L.G.: Technological Innovation as a Process. The Processes of Technological Innovation. Lexington Books (1990). https://www.researchgate.net/publication/291824703_Technological_Innovation_as_a_Process
75. Desouza, K.C., Dawson, G., Chenok, D.: Designing, developing, and deploying artificial intelligence systems: lessons from and for the public sector Bus. Horizons **63** (2), 205–213 (2020)
76. European Parliament: Regulation (EU) 2016/679 of 27 April 2016, General Data Protection Regulation. (2016) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679
77. Pencheva I., Esteve, M., Mikhaylov, S.J: Big data and AI - a transformational shift for government: So, what next for research?. Public Policy Adm. **35**(1), 24−44 (2020) https://doi.org/10.1177/0952076718780537
78. Regan, P.M., Maschino, K.: Artificial intelligence and its impact in public administration. In: Shark, Alan R. (Ed.), National Academy of Public Administration. (2019) https://napawash.org/uploads/Academy_Studies/9781733887106.pdf
79. Gstrein, O. J., Bunnik, A., Zwitter, A.: Ethical, Legal and Social Challenges of Predictive Policing, Catholic Law Review, University Medical Center, Groningen (2019), https://pure.rug.nl/ws/portalfiles/portal/117795047/SSRN_id3447158.pdf
80. Berryhill, J., Heang, K.K., Clogher, R., McBride, K.: Hello, World: Artificial Intelligence and its Use – the Public Sector. OECD Working Papers on Public Governance (2019) https://doi.org/10.1787/19934351

81. Watkins, I.: The utility of level of service inventory - revised (LSI-R) assessments within NSW Correctional environments. (2011). https://www.semanticscholar.org/paper/The-utility-of-level-of-service-inventory-revised-Watkins/550f2eed8f440b4ba0acc459b5558f99086fe666
82. O'Neil, C.: Weapons of Math Destruction. Penguin Books, London (2019)
83. Fry, H.: Hello World: Being Human in the Age of Algorithms Hardcover – September, Transworld Publishers Ltd. (2018)
84. Andrews, D., Bonta, J.: The Psychology of Criminal Conduct. Routledge, 6th edition. (2018)
85. Roth Davies: What is the effect of a defendant's refusal to participate the LSI-R? (2023), https://www.rothdavies.com/criminal-defense/frequently-asked-questions-about-criminal-defense/sentencing-outcomes/what-is-the-effect-of-a-defendants-refusal-to-participate-in-the-lsi-r/
86. Alexandre, A.C.: A Inteligência Artificial é discriminatória? (II Parte), O preconceito racial presente em decisões judiciais coadjuvadas pelos sistemas artificiais "inteligentes", Jornal de Monchique, edition 470, 28th of January (2023)
87. Leslie, D.: Understanding artificial intelligence ethics and safety A guide to AI ethics, including responsible design and implementation of AI systems in the public sector. The Alan Turing Institute, London. (2019). https://www.turing.ac.uk/sites/default/files/2019-08/understanding_artificial_intelligence_ethics_and_safety.pdf
88. Otero, P.: Manual de Direito Administrativo, vol. 1. Editora Almedina, Lisbon (2013)
89. Council of Europe: Ad Hoc Committee on Artificial Intelligence Policy Development Group, Artificial Intelligence in Public Sector (2021). https://rm.coe.int/cahai-pdg-2021-06-2779-3226-6755-v-1/1680a29927
90. Ferrer, X., van Nuenen, T., Such, J.M., Coté, M. Criado, N.: Bias and Discrimination in AI: a cross-disciplinary perspective. IEEE Technol. Soc. Mag. **40**(2), (2021), https://ieeexplore.ieee.org/document/9445793
91. Committee on Standards in Public Life: Artificial Intelligence and Public Standards Life Artificial Intelligence and Public Standards - A Review by the Committee on Standards in Public Life, UK (2020), https://www.assets.publishing.service.gov.uk/government/uploads/system/uploads/attachmentdata/file/868284/Web_Version_AI_and_Public_Standards.PDF
92. Brynjolfsson, E., McAfee, A.: O negócio da inteligência artificial, Inteligência Artificial, Harvard Business Review. Actual Editora, Lisbon (2019)
93. Harrison, T.M., Luna-Reyes, L.F.: Cultivating trustworthy artificial intelligence in digital government. Social Sci. Comput. Rev. **40**(2), 494–511. (2022), https://journals.sagepub.com/doi/abs/10.1177/0894439320980122
94. Olsen, H.P., Slosser, J.L., Hildebrandt, T.T., Wiesener, C.: What's in the Box? The Legal Requirement of Explainability. In: Computationally Aided Decision-Making in Public Administration. Faculty of Law, University of Copenhagen, Elsevier (2019). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402974
95. Janssen, M., Brous, P., Estevez, E., Barbosa, L.S., Janowski, T.: Data governance: Organizing data for trustworthy artificial intelligence Government Information Quarterly, **37**(3), 101493 (2020), https://pure.tudelft.nl/ws/portalfiles/portal/85607215/1_s2.0_S0740624X20302719_main.pdf
96. Whittaker, M., et al.: AI Now Report 2018. AI Now Institute, New York University. (2018). https://ainowinstitute.org/AI_Now_2018_Report.pdf

Author Index

E. González-Esteban et al. (Eds.): The ETHNA System Project, LNCS 13875, p. 249, 2023.
https://doi.org/10.1007/978-3-031-33177-0