Paper for the OECD Blue Sky Forum 2016 – final version, 15 July 2016

Title: Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) – a preliminary framework for RRI dimensions & indicators

Authors: Ingeborg Meijer, Niels Mejlgaard, Ralf Lindner, Richard Woolley, Ismael Rafols, Erich Griesler, Angela Wroblewski, Susanne Buehrer, Jack Stilgoe, Lena Tsipouri, Nikos Maroulis & Viola Peter, representing the MoRRI consortium

Abstract

There has been a growing focus, among both policymakers and researchers, on ideas of responsible research and innovation, as a way to ensure that new sources of public value are captured. The EC defines RRI as "*a process where all societal actors (researchers, citizens, policy makers, business, third sector organisations etc) work together during the whole R&I process in order to better align R&I outcomes to the values, needs and expectations of European society*". In operational terms, it brings together under the RRI umbrella dimensions of engagement, gender, education, open access, ethics and governance. RRI is a highly relevant area for policy. Therefore the EC has commissioned a study on 'Monitoring the Evolution and Benefits of Responsible Research and Innovation' (MoRRI). It contributes conceptual work on RRI, provides extensive exploration of existing metrics capturing RRI, and develops new indicators requiring primary data collection.
Indicator development is based upon an intervention logic that builds mainly on the five dimensions mentioned complemented by the horizontal dimension of governance. Monitoring is suggested to build upon a set of 6x6 indicators that were selected from a broader number of potential indicators that were validated in a multi-level process. Initial data collection was used to testing these indicators, whereby the emphasis is put on using both quantitative and qualitative methods for all areas.

The lack of adequate measures of responsibility in research and innovation hampers the mainstreaming of RRI. While 'pairing' responsibility and scientific excellence is an explicit aim for the RRI agenda, they are in reality often perceived as contradictory demands by the individual scientists or viewed as unequally important concerns by the research performing - and research funding – organisations. Identification of useful indicators and metrics for RRI might contribute to bringing issues of responsibility from a peripheral position and closer to the center of activity.
This paper concludes with a reflection on evidence of 'benefits' or impacts of RRI. Being able to identify such benefits goes beyond the rather simplistic idea of monitoring RRI through indicators. Identifiable benefits will further help in shaping and refining indicators. And finding a common denominator in a broad diversity of benefits helps policy in structuring further mainstreaming of RRI.

1. Introduction

In recent years, the notion of Responsible Research and Innovation (RRI) has emerged in European policy making, as a way to ensure that new sources of public value are captured. From the perspective of the 'Science with and for society' scheme of the European Commission (EC), the purpose of promoting RRI is "to build effective cooperation between science and society, to recruit new talent for science and to pair scientific excellence with social awareness and responsibility".

Building on work by Von Schomberg (2013) and others (Stilgoe et al., 2013; Owen et al., 2013; Pandza and Ellwood, 2013; and Rip 2014), the EC defines RRI as "*a process where all societal actors (researchers, citizens, policy makers, business, third sector organisations etc) work together during the whole R&I process in order to better align R&I outcomes to the values, needs and expectations of European society*" [1]. Conceptually RRI reflects (previous) strands of activities such as anticipatory governance (Karinen and Guston, 2010), Constructive, Real-Time and other forms of technology assessment (Rip et al., 1995; Guston and Sarewitz, 2002; Grin and Grunwald, 2000), upstream engagement (Wilsdon and Willis, 2004), value-sensitive design (Friedman, 1996) socio-technical integration (Fisher et al., 2006), social responsibility (Glerup and Horst, 2014) and transdisciplinarity (Wickson and Carew, 2014).

The RRI definition is firmly anchored in European policy processes and values. Starting from a rather confined academic debate calling for responsible innovation (Hellström, 2013), the idea is now part of the EU's research and innovation policy as a cross-cutting theme in the current framework programme Horizon2020 and the European Union's ambitious goal to become a true Innovation Union, in which research and innovation are the main drivers of competitiveness, jobs, sustainable growth and social progress. At the same time a number of grand societal challenges need to be addressed as well such as climate change or health and wellbeing. Orientating research and innovation towards societal goals is being reinforced by the debate on RRI since RRI emphasizes specific qualities of research and innovation practices, and aims to redefine the roles and responsibilities at science-society interfaces (Nielsen et al., 2015). While RRI is supporting the notion how research and innovation can contribute to our future society, it likewise raises the question how to take into account and mitigate the unanticipated, unintended and undesirable consequences of emerging science and innovation.

In operational terms, however, the EC brings dimensions of public engagement, gender equality, science literacy and science education, open access, ethics and governance under the RRI umbrella. These six RRI 'keys' have informed the composition of the work program for SwafS and the featuring of RRI as a cross-cutting issue of Horizon 2020, intended to be embedded across the priorities of the funding programme.

So far, efforts to 'mainstream' RRI across the European research area have been modestly successful (Mejlgaard and Griessler, 2016). Studies indicate significant obstacles, pertaining not least to disincentivizing reward structures at both organisational and individual level[2]. While 'pairing' responsibility and scientific

[1] https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society.

[2] More information available at http://www.rri-tools.eu/documents/10184/107098/RRITools_D2.2-AnalysisNeeds+ConstraintsStakeholderGroupsRRI.pdf/83c55909-118c-4cad-b7e4-74d5a770c8a1

excellence is an explicit aim for the RRI agenda, they are in reality often perceived as contradictory demands by the individual scientists or viewed as unequally important concerns by the research performing - and research funding - organisations. While the production of high impact publications is, e.g., considered a core academic activity clearly carrying merit, engaging in public outreach or stakeholder dialogues might easily be considered peripheral activities without straightforward value for the individual scientists.

Moreover, the lack of adequate measures of responsibility in research and innovation further hampers the mainstreaming of RRI. Inability to evaluate, compare, and benchmark 'performance' in terms of RRI at the national as well as disaggregated levels, constitutes a barrier to any revision of reward schemes and dilutes the potential vitality of the organisational or national 'horse race' for high performance in this area. Identification of useful indicators and metrics for RRI might then contribute to bringing issues of responsibility from a peripheral position and closer to the center of activity. Finally, evidence of 'benefits' or impacts of RRI are likely to further contribute to mainstreaming of RRI activities at the level of member states and R&I organisations, and possibly also to increased attention to this cross-cutting issue across the H2020 programmes.

Against this backdrop, the EC has commissioned a study on 'Monitoring the Evolution and Benefits of Responsible Research and Innovation' (MoRRI). It contributes conceptual work on RRI, provides extensive exploration of existing metrics capturing RRI, and develops new indicators requiring primary data collection.

In the next paragraphs, this paper elaborates on further conceptually defining what is (and what is not) RRI, in order to lay the foundations for a broad-based agreement for policy intervention; and the key's and areas to be monitored (par 2); the intervention logic for monitoring and evaluation of RRI keys (par. 3); a preliminary selection of indicators across all keys and reflection on data collection (par. 4); and finally a concluding paragraph offering a critical reflection on, indicators and the 'benefits' of RRI and the use of RRI indicators for mainstreaming of RRI (par 5.). While this project puts things in a European perspective, the aim is not losing sight of global evolution. The paper builds upon previous project deliverables (MoRRI, 2015)

2. The Six RRI keys

The first phase of the indicator development work in the project has been carried out and consisted of review of the literature on the six RRI keys. The first step was to identify the relevant literature and documents dealing with RRI issues. A systematic review covering a variety of RRI related documents, including academic literature, EC and other policy documents, conferences and on-going projects, provided central insights into each of the six RRI key dimensions, their policy context and main definitional elements. Informed by these theoretical and conceptual explorations, the main next step has been to develop a functional vocabulary covering each of the six RRI keys. This procedural step has involved a stocktaking and assessment of all existing data sources considered relevant in the monitoring of the six RRI dimensions, including reflections on data gaps and assessments of the need for primary data collection in the subsequent tasks of the project. The six keys are:

2.1 Citizen engagement and participation of societal actors in research and innovation, in short Public Engagement (PE)

The synthesis of core literature shows that 'citizen engagement and participation of societal actors in research and innovation', or PE, is today a rich and diversified field of practice and academic studies, and the concept of PE is multifaceted. The PE field has been reframed and transformed within the last decades and a general turn from one-way and top-down models of communication towards increased focus on 'new' dialogue-based approaches characterizes the development of the field (Bauer et al., 2007).

In 1985, The Royal Society in London issued a report on 'The Public Understanding of Science', which has been an influential document in the modern history of public engagement with science. The main message of the report was that given considerable public investments in science and technology, science should be transparent and scientists should account for the societal consequences of their work. The report highlighted the intimate connection between national prosperity, science, and technological progress, and the pervasiveness of science and technology in the everyday life of citizens. It identified a need to strengthen science communication efforts for two purposes: first, because it basically is not possible to navigate successfully in modern societies without an overall understanding of science. Second, science and technology are fundamental forces in the broader innovation system, which generates progress and prosperity. There is a need to tell this story to the public in order to generate a broader appreciation of science and technology (Wynne, 1995). In this respect, 'understanding' science is not merely a question of being interested and knowledgeable, but rather a question of appreciating and acknowledging the importance of science and technology as main drivers of economic and societal progress.

Building on the results of the literature review, a functional vocabulary of PE was developed that presents the definitions and terminology related to PE that will allow an empirical and practical approach to the concept of PE. From this, it was concluded that there is no singular conception of 'engagement' and no single model of its implementation. Rowe and Frewer (2005) develop a typology of PE mechanisms based on the direction of the flow of information between representatives of the public on the one hand and the sponsors of engagement initiatives on the other hand, resulting in a differentiation between 'public communication', 'public consultation', and 'public participation'. Ravn, Mejlgaard and Rask (2014), informed by an inventorying of 250 specific engagement initiatives across Europe and beyond, add to this PE mechanisms like 'public activism' and 'public deliberation'. Together, in the MoRRI project, the operational understanding of PE while recognizing the complexity of objectives for PE and the variation in mechanisms for engagement, distinguishes five main categories of PE, namely 'public communication', 'public activism', 'public consultation', 'public deliberation', and 'public participation'. This categorisation was originally developed by the PE2020 project[3].

---

[3] More information on the project available at www.pe2020.eu

Public communication – *the aim is to inform and/or educate citizens*. The flow of information constitutes one-way communication from sponsors to public representatives, and no specific mechanisms exist to handle public feedback (examples include public hearings, public meetings and awareness raising activities).
Public activism – *the aim is to inform decision-makers and create awareness to influence decision-making processes.* The information flow is conveyed in one-way communication from citizens to sponsors but not on the initiative of the sponsors, which characterise the 'public consultation' category (examples include demonstrations and protests).
Public consultation – *the aim is to inform decision-makers about public opinions on certain topics.* These opinions are sought from the sponsors of the PE initiative and no dialogue is implemented. Thus, in this case, the one-way communication is conveyed from citizens to sponsors on the initiative of sponsors (examples include citizens' panels, planning for real, focus groups and science shops).
Public deliberation – *the aim is to facilitate group deliberation on policy issues, where the outcome may impact decision-making.* Information is exchanged between sponsors and public representatives and a dialogue is facilitated. The flow of information constitutes two-way communication (examples include 'mini publics' such as consensus conferences, citizen juries, deliberative opinion polling).
Public participation – *the aim is to assign partly or full decision-making-power to citizens on policy issues.* Information is exchanged between sponsors and public representatives and a dialogue is facilitated. The flow of information constitutes two-way communication (examples include co-governance and direct democracy mechanisms such as participatory budgeting, youth councils and binding referendums).

2.2 Science literacy and scientific education (SLSE)

Science literacy and scientific education have been topics of academic and public discussions for a long time and continues to do so, especially in the light of the challenges of modern societies. Given its long history the field is, on one hand, well researched. On the other hand, substantial questions still remain to be answered. The subject of science literacy only became a topic of systematic study in the 1930s when John Dewey argued that "*the responsibility of science cannot be fulfilled by methods that are chiefly concerned with self-perpetuation of specialized science to the neglect of influencing the much larger number to adopt into the very make-up of their minds those attitudes of open-mindedness, intellectual integrity, observation and interest in testing their opinions and beliefs, that are characteristics of the scientific attitude."* (Dewey, 1934)
Over the past five decades, the field of science literacy and science education has experienced at least two major shifts, leading to the current co-existence of three paradigms: the deficit model (i.e. the idea that science is "too complicated" for the general public to understand), the dialogue model (i.e lay people have knowledge and competencies which enhance and complete those of scientists and specialists" (Bucchi, 2008; Callon, 1999) and the participation model (or co-production of knowledge) where the need and right of the public to participate in the discussion has been put to the forefront. The dialogue model developed following conclusions in the influential Royal Society report on Public engagement (1985). The report argued that the deficit of the public was deriving less from knowledge but rather from attitudes. It was claimed that the public's attitude towards science and technology is not sufficiently positive and that dangers existed that citizens would become negative or even espouse anti-science

attitudes (Bauer et al, 2007). In order to recover the deficits, activities in this period focused on changing attitudes and marketing the image of science.

For the purpose of this project we define science literacy as the ability of citizens to read about, comprehend and express opinion about science, as well as the ability to contribute to "doing science". By building on this idea, the focus of our understanding of science literacy is put on the idea of developing capacities for science and innovation. Science literacy is generated through activities aiming to provide citizens with a deeper understanding of science, to shape their attitudes towards science and to develop their abilities to contribute to science and science-related policy-making.

The operational understanding of SLSE applies a tripartite categorisation for the multifaceted field of science literacy. Science literacy can be generated through three main mechanisms:

Science education aims at educating (especially young) citizens about scientific facts (textbook knowledge), the norms of science and the way science is 'done' as well as at conveying a positive 'image' of sciences. However, it also provides the opportunity to reflect and question science and the 'truths' it produces critically. It takes place in institutions in early childhood education and care, the school system, higher education, vocational education and training as well as in lifelong-learning. Science education is the basis for science literacy.

Science communication activities aim at educating citizens of all ages about science as well as at generating awareness of science-related issues and a positive image of/attitude towards science. These activities can take direct forms (for instance through open days, museums or science centres) or be more indirect with mediators between the scientists and the public (e.g. via science journalists and their products such as TV programmes or media articles etc). Generally, a large number of different institutions are involved in science communication.

Co-production of knowledge is characterised by a co-creation of knowledge through cooperation of scientific experts and non-experts. Including the co-production of knowledge in the dimension of SLSE alters the way we think about the public and its role in science and innovation, from a mere receiver and customer to an active agent of change. Citizens co-produce scientific data, possibly help in their interpretation and analysis and frame research questions. The aspect of *co-production of knowledge* is clearly interlinked with mechanisms and activities carried out within the field of public engagement.

2.3 Gender equality (GE)

Since the turn of the century, the topic of gender equality in science and research has been intensively discussed. Accordingly, a broad range of literature, pilot projects and empirical evidence is available which deals with gender inequalities in this area. This provides the starting point for the discussion on gender within responsible research and innovation (RRI) and the development of indicators for the gender dimension in RRI.

Since the early 1990s, the presence of women in science has gained increasing interest in political as well as scientific debate. This debate was initially supported by calls for social justice and was embedded in the development of general antidiscrimination policies at both national and European level aimed at establishing equal rights for women in employment. Corresponding research focused on the career paths of men and women as well as on the complex interplay between the institutional arrangements and personal preferences that might serve to explain the underrepresentation of women, especially at the top levels (Caprile et al., 2012). Since the turn of the century, economic arguments have also been used increasingly to justify gender equality policies:

- In the European Commission's (EC) view, realizing Europe's ambition to achieve a *competitive knowledge-based society* will require an increase in the number of researchers (European Commission: The Wake-Up Call for European Industry 2003). In order to achieve goals like 'competitiveness', 'innovation' and a 'knowledge-based society', it is evident that the talents and potential of women have to be developed, mobilized, leveraged and used more actively, deeply and completely.
- From the science and technology perspective, *'gendered innovations*' enhance excellence in science, medicine and engineering both in terms of knowledge and personnel. 'Gendered innovation' is defined as the process that integrates sex and gender analysis into all phases of basic and applied research to assure excellence and quality in outcomes (Schiebinger and Schraudner, 2011).
- In the business sector's view, the reasons why *gender diversity* should be taken into serious consideration lie in women's talents, the economic power of women, the changing market structure and the positive impact of women on organizational excellence and financial performance (Catalyst, 2004; McKinsey, 2007).

The issue of the under-representation of women in top positions both in academia and in the business sector is widely discussed. The literature interprets and explains this persistent gender segregation at three levels: the individual level, the institutional level and the social/cultural level. In 2007, the European Commission changed its policy approach from 'fixing the women' to 'fixing the institutions' in line with the process related approach of gender mainstreaming (Lipinsky, 2014). As a consequence of all the above, a broad policy mix has been developed to support women and overcome gendered structures. This includes both the provision of career support for women as well as institutional measures.

Following the recent political and scientific discourse, the operational understanding of gender equality is defined as a three-dimensional construct addressing 1) the (under-) representation of women in research and innovation with the objective to reduce gender segregation, 2) the structural and organisational changes in research institutions with the aim to break down structural gender barriers by means of action plans, gender budgeting, among others actions, and 3) the inclusion of gender in R&I content, thereby aiming at:

- **Integration of women in all fields and at all levels** in research and innovation (reduction of horizontal and vertical segregation); This comprises measures to promote women in fields, where they are underrepresented as well as to increase female participation in management and decision-making positions. The goal here is to reduce gender segregation.
- **Structural change in research institutions** in order to abolish structural barriers for women. This comprises structural measures aimed at revising existing organisational arrangements to progressively eliminate barriers impeding women's advancement to top positions and factors inducing women to drop out of science.
- **Integration of gender in the content of research and innovation** to ensure that women's needs and interests are adequately addressed. Explicit gender issues are rarely included in the content, and it is argued that research results are not valid or

reliable if they only consider male research subjects. This is legitimised by the gender mainstreaming strategy on the one hand and by quality standards in science and research on the other (Caprile et al., 2012).

The gender equality key is closely connected with the ethics and governance dimension.

2.4 Open Access (OA)

Historically, open science relates to the need to build a publicly recognised reputation. The scholarly tradition of open knowledge was turned into a procedure for establishing knowledge claims that could be evaluated and recognised by peers and then utilized by others. Knowledge was considered a public good, and likewise a publication (any kind) as well. Since then, propertisation of knowledge occurred through copyright imposed by the academic publishing market. A lack of policy coordination and/or framework conditions impeded the free movement of research activities and knowledge, hindering access to publicly funded research results and knowledge transfer. A diversity of national policies, legal requirements and practices regarding knowledge transfer as well as open access to scientific publications and scientific data adversely affected the wider dissemination, access to and use of knowledge created with public funds. Open, however, should apply to all components of the research process, and not be restricted to the outcomes only. Open will then need to be embedded in the research process from start to finish. Such changes may impact the entire research cycle and its organisation, from the inception of research to its publication. Research organisations, funding bodies, career paths of researchers are affected as well as the science system, e.g. the rise of new scientific disciplines, new pathways in publishing, and different scientific reputation systems, and how the quality and impact of research is evaluated.

Open access (OA) is the idea of making research results freely available to anyone who wants to access and re-use them. Nowadays, it is widely propagated that making research results more accessible contributes to better and more efficient science, and to innovation in the public and private sectors (EC 2012, 2014). In 2003 *the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* signaled the start of open access policies, followed by the *Open Access pilot initiative* within the mainstream Framework Programme itself (in addition to the ERC) which was launched by the EC in 2008. Subsequently, the policy movement to enhance Open Access to research outputs led to the launch of a series of projects; one of them, funded under the EC Research Infrastructures programme, is the OpenAIRE pilot action[4]. In 2012, the European Commission encouraged all EU Member States to put public-funded research results in the public sphere in order to make science better and strengthen their knowledge-based economy. As other challenges need to be addressed such as infrastructure, intellectual property rights, content-mining and alternative metrics, but also inter-institutional, inter-disciplinary and international collaboration among all actors in research and innovation, the European Commission is now moving decisively from 'Open access' into the broader picture of 'Open science' in order to "*enhance open circulation of knowledge across national borders, including knowledge transfer*". The assumed benefits of Open Science will be more transparency, openness, and networked collaboration. In the long term, European policy hopes to may make science more

[4] More information on the project available at https://www.openaire.eu/

efficient, reliable and responsive to the European Grand Challenges of our times as well as foster Open innovation.

The operational understanding of OA in the context of the MoRRI project is reviewed from three separate angles: Starting from the general policy concept of open science that focuses on a lack of policy coordination and/or framework conditions impeding the free movement of research activities and knowledge, hindering access to publicly funded research results and knowledge transfer (van den Eynden and Bishop, 2014), two specific aspects are of interest: 'The Open Access instrument for publications' and 'Developments in Open data'.

- **Open Access publications:** The term Open Access originally referred to the provision of free access to peer-reviewed academic publications. The 'Assessment of Open Access publishing' is complicated by the growing diversity of what counts as Open Access, the copyright restrictions for when a publication can be made openly accessible, and the lack of clear and consistent identification of Open Access publications in bibliographic data. There are two formal operational paths to access through Open Access journals and self-archiving in repositories, subsequently referred to as *Gold Open Access* (open access journals), and *Green Open Access* (also called 'self-archiving'). The important and on-going debate in the assessment of OA is whether OA publishing yields increased citation impact. While there are conflicting reports in the debate about an Open Access citation advantage (OACA), heightened attention to this issue has increased our understanding about Open Access publishing more generally (Björk et al., 2009; Archambault et al., 2014).
- **Open Research Data (OD)** is a relatively new and emerging field of scholarship, and systematized data sources are still fairly scarce compared to the data availability on issues related to *Open Access* (Costas et al., 2013; Farhan et al., 2013). Data and datasets are central for empirically oriented science and scholarship. They can be very diverse (e.g. archaeological, biological, genetic, economic, mathematical, astronomic, etc.) and once collected, the same data can be used by a variety of researchers from different institutes, disciplines and organisations to produce new results. Open data and data sharing have proven to create many benefits, such as stronger open science; a higher efficiency in the use (and reuse) of scientific resources; the possibilities of expanding new research lines. Despite these main benefits, there are cultural (e.g. trust and control) and technical (e.g. standards and metadata) issues that hamper widespread data curation, sharing and use.

The other big challenge for OA is re-uniting publications with public funding and to interlink publications with research data. Such improved interlinking would allow for the investigation of research results, beyond the limits of project-based funding, and provide the data needed for the exploration of longer-term results of public research funding.

2.5 Ethics

The term "ethics" is rarely defined comprehensively; it seems that most authors assume that the reader already knows what ethics in the context of research and innovation is about (Brom et al., 2015). One example of a definition is *"A common platform for deliberation and discussion of values in society, that is based on perceptions of right and wrong, is influenced by cultural norms, and aims at informing policy making"* (Ladikas et al., 2015). Others make a distinction between ethics as a scientific discipline and moral.

Griessler and Littig (2006) follow the literature in stating that: “Ethics as a scientific discipline is concerned with normative rules for everybody, which other than moral should be used to evaluate and not to guide actions” Schicktanz et al. (2012) define morality as “a set of rules and values actually guiding individual life and social interaction – and ethics as its normative reflection, justification or critique in view of validity, desirability and legitimacy”. Ethics is discussed in different political contexts (global, European, national), reflecting the institutionalization of ethics in science and innovation. But ethics can be discussed at the informal level as well (Felt et al., 2009), e.g. by engaging experts and laypeople in a symmetrical discussion of the ethics of particular biomedical fields of research.
Key qualities of ethics in the context of R&I demanded by many authors are process qualities such as:
1. Openness towards stakeholders and the public;
2. Public participation (including information, consultation of, and with deliberation public);
3. Transparency and accountability of processes;
4. Thematic openness in terms of which questions can be raised;
5. Systematic argumentation in terms of a priority of arguing over (political) bargaining (this also includes scholarly integrity).

The operational understanding of ethics in research and innovation and delineation of the institutionalization of ethics is categorized in: ‘ethical governance’, ‘ethical deliberation’, and ‘ethical reflection’.
Ethical governance: I.e. “institutionalising ethics debate in terms of the implementation of standards in research ethics in science, technology and innovation policies” (Brom et al. 2015). In the area of ethical governance, ethics commissions are a major governance instrument, advising e.g. national government on policy making. Another important instrument to govern ethics in science and technology are ethical codes and soft law.
Ethical deliberation: I.e. “institutionalising ethics debate that raise issues in science and technological developments in science, technology and innovation policies”. One important instrument for ethical deliberation is technology assessment (TA).
Ethical reflection: I.e. “institutionalising ethics debate that support critical reflection and engagement in debates on research standards, emerging technology issues and social justice in science, technology and innovation policies”. Instead of focusing on advisory committees ethical reflection is advocated through *round table formats* to bring together experts and laypeople.

2.6 Governance

The relationship between governance, research and innovation is far from simple, and far from linear. Not only are science and innovation governed in various ways, some of which may be considered more responsible than others, but also science and innovation are a vital and increasing part of our governance regimes. ‘Governance’ here refers to control or management. It can be found not just in the state, but also in businesses or any social organization. In this sense, governance goes substantially beyond ‘government’. Here we take a definition from Borrás and Edler (2014), who define governance in relation to STI as the *"[...] way in which societal and state actors intentionally interact in order to transform ST&I systems, by regulating issues of societal concern, defining processes and direction of how technological artefacts and innovations* *are produced, and shaping how these are introduced, absorbed, diffused and used within society and economy."* Governance involves the establishment of goals, setting up of means, and verification of performance. This includes providing, distributing, and regulating. When applied to science, governance is often interpreted to mean 'regulation', suggesting a restriction of freedom. Regulation is a subset of governance focused on steering and confirming trajectories, as opposed to providing and distributing.

The governance of science is concerned with how knowledge is produced and how it is distributed. The governance of innovation is far broader. The concern with the latter goes beyond concerns about technology and its regulation. Scientists govern how knowledge is produced, certified, made credible and communicated. RRI, then, is clearly aligned with particular ideas of governance and a set of innovations towards this end, such as public deliberation, lay membership of expert committees, transparency and multidisciplinary collaboration. In particular, RRI governance should be seen in the light of recent moves and frameworks aiming towards 'anticipatory governance' (Barben et al., 2008; Guston, 2014).

The operational understanding of the horizontal dimension of governance is defined as steering innovation through the establishment of goals, the establishment of means and the verification of performance. Particularly important in the case of the governance of science is the realisation that much governance happens within and is done by the scientific community itself.

Below we categorise a typology of governance approaches that helps to structure discussions about changing governance (adapted from Hagendijk and Irwin, 2006):

Discretionary governance: Policies in this category are made without explicit interaction with 'the public'. Governance is presented primarily as a matter for government, which is seen as serving universal goals of progress.

Corporatist governance: This involves a formal recognition of differences of interest as an input to negotiation. As negotiation takes place within a closed or highly regulated space, the decisive feature of this mode is the admission of stakeholders.

Educational governance: This assumes that policies for science and technology have foundered on the shoals of public ignorance. Hence, it is necessary to create an informed citizenry.

Market governance: Science and technology are best regulated by demand and supply. The value of science comes from the surplus value created through its commercialization and contribution to the generation of wealth. The public participates as customers and consumers.

Agonistic governance: This form of governance occurs in a context of confrontation and adversity. The storage of nuclear waste in the UK is a case where policy seems to have stalled in the face of public opposition: opposition to GM foods has also taken agonistic form.

Deliberative governance: This rests on the assumption that open debate and engagement can create a satisfactory foundation for decision-making. In this mode, the public are not consumers of science, but rather 'scientific citizens'.

3. The intervention logic for monitoring and evaluation of RRI keys

Indicator development is based upon an intervention logic that builds mainly on the six keys mentioned. Monitoring is suggested to build upon a set of appropriate indicators that are selected from a broader number of potential indicators that were validated in a multi-level process. Initial data collection is used to testing these indicators, whereby the emphasis is put on using both quantitative and qualitative methods for all areas. The first phase of the indicator development work in the project has been carried out and consisted of:

- Assessments of finished and ongoing (inter)national and EC projects to make an inventory on used and potential indicators for the six keys;
- The extent of data availability within separate key categorisations/typologies as described above was assessed through a cross-read of data explorations vis-à-vis functional vocabularies

The literature review together with the assessment of projects and data availability have resulted in a tentative list of altogether 98 existing indicators / data sources that were considered potentially useful in measuring and capturing core aspects of RRI.

In the next step, these promising 98 existing indicators and data sources were classified and assessed with respect to their *relevance/proximity* and *robustness/quality* as individual measures of RRI at specified keys and levels of analysis. Moreover, an aggregate assessment and classification of the overall *relevance/proximity*, *robustness/quality*, and *data richness* of the available indicators have been conducted as well.

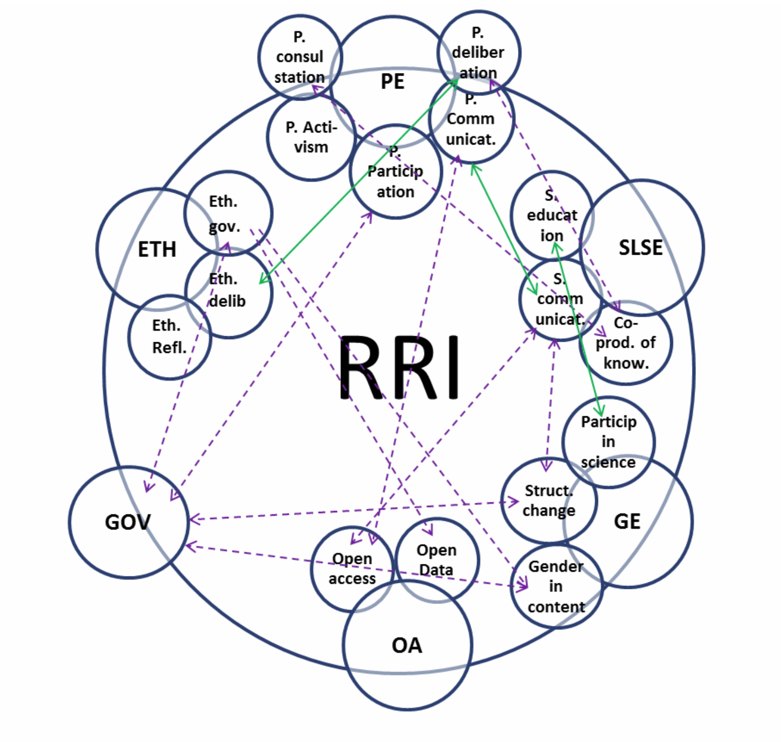
To develop a solid conceptual framework capable of addressing the complex nature of RRI in the best possible manner, the MoRRI project borrows ideas from the evaluation literature. It introduces the 'intervention logic model' as a starting point for monitoring various types of impacts and benefits of RRI. The intervention logic model is based on the explanatory idea that complex policy problems are characterised by a series of issues or problems that need to be addressed, a set of *inputs* which are applied to a series of activities, which generate *outputs* which in turn lead to *outcomes* or the resolution of the problems. This logic informs and relates to a 'theory of change', i.e. an assumption or hypothesis of why an intervention will succeed in producing desired outcomes and impact(s). A 'theory of change' specifies how activities are expected to lead to interim and longer term outcomes. The elaboration of the theory of change can strengthen the case for attributing observed changes to an action and is thus an important aspect of the method of approach to the evaluation. An important aspect of the logic model is the identification and description of *key contextual, external* factors that could influence the intervention either positively or negatively.

The main characteristics of the elements of the intervention logic model applied in the MoRRI project can thus be summarised as follows:

- *Context indicators* provide information on the environment and overall situation in a country and across countries;
- *Input indicators* concern the activities carried out, measures taken, structures created and resources allocated to promote RRI. This type of indicators focuses on the system and consistency of the RRI related initiatives.
- *Output indicators* address the immediate and direct results of these activities, while indicators of *outcome* scrutinize the more far-reaching and long-term achievements and perceived benefits of the RRI work.

A main conclusion derived from the synthesis and assessment of the 98 promising RRI indicators, is that the *input* and *output* dimensions of the intervention logic model comprise the most comprehensive and saturated indicators of RRI. Particularly the RRI dimensions of Public Engagement and Gender Equality contain highly relevant, robust and data rich input level indicators, whereas the available Open Access and Ethics indicators, while relevant, are more difficult to assess in terms of robustness and richness of data. Moreover, a data gap exists for the Science Literacy and Science Education dimension at the input level.

Figure 1: Existing and potential interlinkages/overlaps *between* RRI (sub-)categories (Source: MoRRI deliverable 3.1)



Interlinkages among RRI (sub-)keys/categories

While the six keys individually provided a number of potential indicators, the keys are not independent of each other. As described above, interlinkages exist among and between the conceptual categories of each RRI key. In order to account for the analysis of such interlinkages, the following questions were addressed:

- Are the analytical sub-categories within each RRI key sufficiently covered?
- When considering internal overlaps among sub-categories, and the relevance of each sub-categories, do all categories need to be equally well covered?
- Do sub-categories need to be equally well-represented at all levels in the intervention logic model and across aggregation levels?
- To which extent do RRI keys and their respective sub-dimensions intersect, and how do such intersections influence data coverage?

To further guide the collection of primary and secondary data, while ensuring consistency across the six RRI keys at different intervention- and data levels, existing and potential interlinkages between RRI dimensions and sub-dimensions are depicted in figure 1. The coloured arrows reflect whether interlinkages are directly addressed (green) in the existing set of indicators, whereas the purple arrows display potential interlinkages that could be further explored. The direction of the arrows indicates whether the interrelations are reciprocal or nonreciprocal. Several of the indicators addressing the *Public Communication* aspect of PE, for instance, bear clear relevance to, and reveal actual overlaps with, the SLSE sub-category of *science communication*. Similarly, a range of other potential interlinkages could fruitfully be further explicated and explored when identifying and establishing the final set of indicators.

4. A preliminary selection of indicators across all keys and reflection on data collection

The moderated list of indicators and the related assessment process gave rise to cardinal questions concerning the adequacy and accuracy of the individual and compiled set of indicators in covering the general concept of RRI and its six keys. In a joint workshop in 2015 among the project partners, the main criteria for selecting indicators were discussed as well as the following issues:
[a] does the moderated list of indicators adequately cover potential indicator candidates for sustained data collection (i.e. time series data)?
[b] Will it be possible to compile a set of data collection methods through which several metrics across keys can be harvested?
[c] Does the moderated list of indicators fully capture EU28 while also adequately covering the concept of RRI?
The joint deliberations resulted in a common understanding of the framework for the final indicator design, taking into account the view of a number of EC recommendations. The main criteria to guide the identification and construction of the final set of indicators are:

• **EU 28 coverage**: Ideally, the set of RRI indicators should adequately cover EU28 for the sake of representativeness. In a policy perspective, the comparative element also remains significant as a means to gain insights into country performance vis-à-vis responsible research and innovation.
• **Possibility for sustained data collection**: Indicators that could be candidates for long-term, repetitive data collection (beyond the life span of the project) should be pursued and taken into consideration in the design of the primary data collection. A few indicators are not considered well suited for sustained data collection. In most cases, however, these indicators are based on secondary data, which are fairly easy to harvest.
• **RRI conceptual coverage**: The indicators should adequately cover all six keys. Ideally, a number of indicators should be 'multi-dimensional' in character, since such indicators are considered better at capturing the core RRI notion. It should be noted that it will be empirically tested how indicators cluster together once the data is available.
• **Focus on input and output indicators:** Based on the assessment of existing indicators/secondary data, it is recommended that MoRRI should focus on the actually implemented RRI activities (inputs) as well as the outputs generated (output). In line with the RRI expert group (EC 2015), the argument is supported of maintaining a focus on the 'short and midterm effects', while underlining that 'impact evaluation is shifting from (end)product to process, and from verdicts/judgments to learning and improving'.

• **Balance between targeted social actors**: Like the different RRI keys need to be adequately represented in the final set of indicators, so do the various societal actors and stakeholders, both individually and collectively. For instance, if the project is to reach a set of reasonably comprehensive RRI indicators, key social actors representing civil society, research communities, governmental institutions, business and industry, among others, need to be addressed with regard to their various roles in research and innovation processes.
• **Qualitative and quantitative data**: Considering the uncharted field of responsible research and innovation, and consequently the dearth of comprehensive sets of RRI indicators, qualitative data enabling in-depth understandings of the complex and multi-dimensional RRI field are seen as vital in the identification and construction of indicators. Such approaches do however 'rarely provide data that are straight-forwardly applicable in terms of benchmarking, and for the purposes of the monitoring study, it will be necessary to translate qualitative material into 'quantitative' indicators and measures that allow for comparisons across countries' (Davis et al., 2012).
• **General quality assessment criteria:** The criteria of *robustness/quality* concern the validity and reliability of indicators in measuring specified keys and analytical levels of RRI. The issues of validity and reliability do however also constitute more generic quality criteria that need to be taken into consideration in the construction of RRI indicators in general. The most important aspects stated below:
- *Content validity* concerns the extent to which the content or theoretical construct of the indicator matches the content domain it has been defined to measure (Hertog et al. 2014). The issue of indicator attribution constitutes a crucial element in ensuring the content validity of the identified measures.
- *Reliability* concerns the quality, consistency and comparability of the underlying data forming the basis of the identified indicators. As mentioned by Den Hertog et al. (2014), one major issue when drawing on secondary data is that "each country uses its own specific data sets, measurement methods, and definitions. Although supranational organisations such as the OECD and EU have made great progress in unifying international data collection, substantial differences exist between countries. The basic problem is that it is sometimes difficult to tell whether (or to what extent) the differences in a model between countries are real or rather constructs due to differences in measurements". In other words, it seems crucial to account for the actual consistency and comparability of the available aggregate data derived from country-specific data-sets.
- *Indicator coverage bias* aims to clarify whether a bias exists in the structure of the data itself (Den Hertog et al., 2014). The lack of coverage of the humanities and parts of the social sciences in Thomson Reuters' Web of Science, may for instance lead to structural bias in the otherwise highly relevant measures of developments in Open Access publications across scientific disciplines and countries.
- *External validity* addresses the extent to which the data collected on the basis of the indicators are providing information that is generalizable to a broader population of cases, situations or people.

The list of 98 promising indicators outlined for the keys was assessed in terms of robustness, richness and RRI relevance – the latter as defined within the context of MoRRI, with the aim of further delineation and selection of a limited set of indicators. To this end, the following considerations applied:

- Data quality, coherence and availability were considered according to the intervention logic model as well as the levels of aggregation of these promising indicators. More specifically, data coverage was considered in terms of the availability of data across dimensions relating to their characteristics in terms of the intervention logic model, i.e. data describing the context, input, output, and outcome. Context relates to the environment and overall situation in a country; input to the activities carried out, measures taken, structures created or resources provided to address what is done in order to address issues of RRI and whether it is done in a systematic manner; outputs to the immediate or direct results of activities and outcomes relate to the achievements.
  Additionally, availability of data is described according to the level of aggregation of these data, distinguishing data that describe the global level, the national level, the regional level, the institutional level, the programme/project level and the individual level.
- A discussion of data gaps and required data collection at different intervention and data levels to guide and direct the primary data specification;
- In terms of the pillars of RRI, the key of governance – while still constituting a separate key – also functions as an overarching dimension or 'umbrella' concept for the remaining keys In this regard, a great number of indicators identified within the other five keys relate directly to the governance of research and innovation, further indicating that this key can be treated 'as an overarching consideration across the other keys of responsible research and innovation'. (European Union 2012).

The considerations in terms of robustness, richness and RRI relevance, and with regard to EU28 coverage and potential for sustained data collection, resulted in the construction, identification, and specification of relevant metrics and indicators to be used in the subsequent RRI monitoring. Based upon the previous synthesis and assessment of existing indicators and secondary data, we have [a] pinned down the final, core set to 6 indicators for each of the 6 keys of RRI, [b] provided as detailed descriptions of each indicator as possible, and [c] specified the primary and secondary data and procedures guiding the data-collection. It should be noted that the indicators selected tap into RRI practices and could potentially be used for monitoring the evolution of RRI across Europe.
The 36 indicators are listed in Table 1 below.

Table 1: Indicators capturing aspects of RRI

Dimension	Indicator	Indicator Title
Gender equality	GE1	Share of RPOs with gender equality plans
	GE2	Share of female researchers by sector (secondary data)
	GE3	Share of RFOs promoting gender content in research
	GE4	Dissimilarity Index (secondary data)
	GE5	Share of RPOs with policies to promote gender in research content
	GE6	Glass Ceiling Index (secondary data)
	GE7	Gender Pay Gap (secondary data)
	GE8	Share of female heads of RPOs
	GE9	Share of gender-balanced recruitment committees at RPOs
	GE10	Number and share of female inventors and authors
Science literacy and education	SLSE1	Science curricula
	SLSE2	RRI related training
	SLSE3	Science communication (secondary data)
	SLSE4	Citizen science
Ethics	E1	Ethics at the level of Universities
	E2	National Ethics Committees (secondary data)
	E3	Research Funding Organizations Ethics Index
Public	PE1	Models of public involvement in S&T decision making (secondary data)

engagement	PE2	Policy-oriented engagement with science (secondary data)
	PE3	Citizen preferences for active participation in S&T decision making (secondary data)
	PE4	Active information search about controversial technology (secondary data)
	PE5	Public engagement performance mechanisms at the level of research institutions
	PE6	Dedicated resources for PE
	PE7	Embedment of PE activities in the funding structure of key public research funding agencies
	PE8	Public engagement elements as evaluative criteria in research proposal evaluations
	PE9	R&I democratization index
	PE10	National infrastructure for involvement of citizens and societal actors in research and innovation
Open access	OA1	Open Access Literature
	OA2	Data publications and citations per country
	OA3	Social media outreach
	OA4	Public perception of Open Access (secondary data)
	OA5	Funder Mandates (secondary data)
	OA6	RPO support structures for researchers as regards incentives and barriers for data sharing
Governance	GOV1	Composite indicator: Governance for responsible research and innovation (secondary data)
	GOV2	Existence of formal governance structures for RRI within RF and RP organisations
	GOV3	Share of research funding and performing organisations promoting RRI

Out of the 36 selected indicators, 13 exploit secondary data (marked in red in the table) while the remaining 23 require primary data collection. Multiple methods will be applied for this purpose. Table 2 provides an overview.

Table 2: Methods for primary data collection

Methods for collecting primary data	Indicators
'Science in Society' actor survey	PE9, PE10
Research performing organisations survey	GE1, GE5, GE8, GE9, SLSE2, SLSE 4, PE5, PE6, OA6, E1, GOV2, GOV3
Research funding organisations survey	GE3, PE7, PE8, E3, GOV2, GOV3
Register-based data	GE10, OA1, OA2, OA3
Qualitative data, interviews/ desk-research	SLSE1

5. A critical reflection on the 'benefits' of RRI and the use of RRI indicators for mainstreaming RRI

As described in 'the Metric Tide report (Wilsdon et al., 2015), there are growing pressures for accountability and evaluation of public spending on higher education and research, such as demands by policymakers for more strategic intelligence on research quality and impact, and the need for institutions to manage and develop their strategies for research. Against this backdrop, the development, description, and operationalization of 'metrics' remains contested and open to misunderstandings, especially when these metrics are about concepts that are complex, such as RRI, or societal impact in general. Although wider use of (quantitative) indicators could support the transition to a more open, accountable and outward-facing research system, inappropriate indicators could easily create perverse incentives. Hence, the consequences of an RRI Scoreboard with a narrow set of RRI indicators need to be identified,

The rise in production and use of indicators in global governance has not been accompanied by systematic study of and reflection on the implications, possibilities and pitfalls of this practice. There is no agreed meaning of "indicator," but Davis et al., (2012) define indicators as 'a named collection of rank-ordered data that purports to represent the past or projected performance of different units. The data are generated through a process that simplifies raw data about a complex social phenomenon. The data, in this simplified and processed form, are capable of being used to compare particular units of analysis (such as countries or institutions or corporations), synchronically or over time, and to evaluate their performance by reference to one or more standards'.

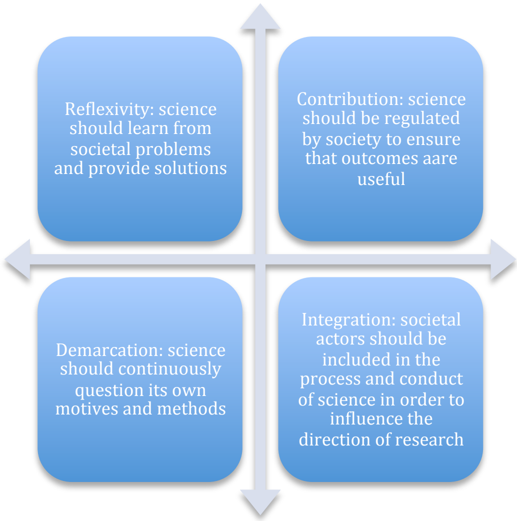
Indicators often take the form of, or can be transformed into, numerical data. A key challenge is whether and how indicators ought to be distinguished from other compilations of numerically rendered data. The differences lie in how indicators simplify "raw" data and then name the resulting product. That simplification can involve aggregation of data from multiple sources, or filtering that excludes certain data, including for instance outliers. They simplify life, and when comparing different value systems one should not compare apples and oranges. The use of any indicator is fine as long as we also discuss what it does not cover, what it may hide or not consider, or what other objections could be there to make clear that indicators does not capture the reality of RRI.
In essence indicators are the ultimate discrepancy between the 'systems world' and 'lifeworld' (Habermas, 1987). The world of science has its own rules of engagement where scientific publications and acquiring funding are crucial 'currencies' (Polanyi, 2000). Criteria and indicators for e.g. engagement with society that are imposed by government for reasons of accountability or steering do not play such a role in the 'lifeworld' of science. In accordance with Habermas (1987), the 'system world' is trying (or even necessary) to govern the science system. In order to act in the academic world, government intends to be accountable for the organisation and execution of science. The information that is necessary to fulfill such task requires structured information gathering and criteria (or indicators). In turn this leads to a dynamic of more specific information along different (RRI) keys that collide with the life world of science. We have to be aware that we do not invade the actual life world of science and innovation in such a way that RRI indicators become counterproductive and create unintended and perverse effects.

The idea of monitoring RRI (through the RRI indicators) are mechanisms from the systems world as. If we want to mainstream RRI, we have to make sure that the lifeworld of science accepts RRI in the same way as they accept the actual currencies of science. From the visioning workshop that was organized by the MoRRI consortium with a number of stakeholders in 2015, a future RRI reality has been described in Lindner et al. (2015). It states that:
- RRI is in your DNA, embedded in daily activities across all actors
- There is a multiple and diverse understanding of excellence in RRI
- There is a merits and incentive structure to support RRI at all levels
- In all steps of the research process – from research agenda setting to evaluation – society is actively involved in the implementation
- And finally, RRI is a creative activity/opportunity instead of a burden

What this vision actually shows is that RRI is not yet mainstream in the life world of science!

How then could we think of further mainstreaming of RRI policies and practices? What are suitable models that describe how the life world of science could interact with the wider society? The RRI policy development aligns well with the debate about social responsibility of science, i.e. for the use of the outcomes of their science (Shapin, 2008). Gelrup and Horst (2014) conceptualize different forms of rationalities in science. They identify four kinds of responsibilities (see figure 2) along two dimensions: The first dimension describes whether the regulation of science is internal or external, and the second dimension describes whether issues of responsibility relate to the process or the outcome of science. The Demarcation and Reflexivity dimensions represent internal regulation of science; the Contribution and Integration rationalities represent the external regulation of science, but either aim for outcomes that can work as solutions (contribution) or to organizing the process (integration).

Figure 2. Overview of four political rationalities (source: Gelrup and Horst, 2014)



(1) Demarcation rationality: departs from a pride in sciences' endeavors, and consider it an honorable profession, which is increasingly suffering from fraud and misconduct. This rationality articulates science as a profession that should have a high autonomy from other actors: outsiders to the scientific realm should not interfere with the discussions about scientists' responsibility.
(2) Reflexivity rationality: acknowledges that science has solved many problems for society but worries that scientist do not assume responsibility for the misapplications of science. Scientists need to make an effort to foresee how their research may affect their surroundings. And the reflexivity rationality describes various techniques to help scientists get better at reflecting on their own practices.
(3) Contribution rationality: articulates science as a social institution; it is part of society and serves certain societal goals. A vision of what is good for society is inherent in these specific goals that science pursues, and likewise society has a decisive role in shaping these visions and goals. According to the contribution rationality, the problem with responsibility in science is that scientists do not see themselves as 'working for the public good', and hence this rationality proposes to enhance outside (societal) control over scientists, in order to make sure that scientists take responsibility to deliver results that are needed by society.
(4) Integration rationality: articulates that science is supposed to be firmly rooted within society, but it is centered around the vision that actors from science and society need to work together as equal partners in order to produce better results. There is a lack of integration between science and other actors in society, and therefore a dialogue between scientists and other actors should be enhanced in order to develop a new kind of 'integrative responsibility' that is suitable for the current specialized societies. This rationality sees scientists as a special kind of citizen that possesses specialized knowledge that can be used to develop society in better directions.

The RRI paradigm in Europe aligns with the Integration rationality in that it stresses that research and innovation activities need to become more responsive to societal challenges and concerns. The indicators that are suggested for the monitoring of RRI in EC countries assume such rationality, but we should be aware that the other rationalities are equally present (cf. Randles et al., 2016) and hence we are pursuing measurement of only part of the lifeworld of science. In fact, the results from the MoRRI visioning workshop show that the Integration rationality is still in the vision stage, and needs further support.

If Open Science in the context of RRI is the main aim, the systems world needs to support science and society developing in the direction of Integration. And in fact, the Responsibility Navigator (Kuhlmann et al., 2016), developed in the Res-AGorA project,[5] supports decision-makers to govern such activities towards more conscious responsibility. What is considered "responsible" will always be defined differently by different actor groups in research, innovation, and society – the Responsibility Navigator is designed to facilitate related debate, negotiation and learning in a constructive and productive way. The Responsibility Navigator supports the identification, development and implementation of measures and procedures that can transform research and innovation in such a way that responsibility becomes an institutionalised ambition.

In the MoRRI project we developed the set of indicators to monitor the progress along the RRI keys. Subsequently we reflected on the mainstreaming of RRI in the life world of science and showed that this is a vision rather than a reality. However, the actual goal is to generate insights into the benefits emerging from RRI: what type of benefits can be expected? Obviously, it is much more challenging to monitor benefits because, when emerging from scientific discoveries, they are almost always associated with risks and uncertainties that take time to measure/eliminate. In the MoRRI project a set of case studies was carried out to obtain an integrative and broad view of how each of the individual keys relates to each other and to identify – until now – unknown benefits. Being able to identify such benefits goes beyond the rather simplistic idea of monitoring RRI through indicators. Identifiable benefits will further help in shaping and refining indicators. Benefits do not arise in isolation: the pathways that lead to benefit are as diverse as the number of case studies. Finding a common denominator in such diversity helps in structuring the benefits.

In the MoRRI project we describe pathways to benefit as having a causal relationship with observable 'first' and 'second order' effects. For instance, gender policy interventions (cause) focus on the right for equal representation, a gendered content of science and the addressing of barriers that prohibit these two elements. The ultimate first order effect of such policy is a research and innovation system free of gender bias. The second order effects of such policy are democratic (equal representation at all levels), economic (better designed products), and social (more inclusive) benefits to society broadly and to the world beyond. Impact pathways lead to cognitive transformations, which are related to the content of science; procedural transformations which are related to the enactment of the science system; and competence and capability transformations which effect all actors in the R&I system (and beyond).

[5] More information on the project available at www.res-agora.eu

Pathways to benefit are always context-dependent, however it is increasingly apparent that RRI drives a number of underlying processes leading toward more 'responsible' configurations of science in and for society. Indeed, many case studies demonstrate processes of pluralization, legitimization and inclusion which work in the direction of producing both the direct benefit of a recognizably more responsible R&I system and a range of more diffused democratic, economic and societal benefits for citizens and stakeholders at large..

Disclaimer:
The MoRRI consortium will use the STI 2016 Conference in Valencia as a platform for presenting and discussing the compilation of indicators tapping into the notion of RRI, addressing specifically the peripheral status of RRI compared to other activities in research and innovation; both in the sense of inadequate measures and in the sense of the lower status of this domain of activity. Furthermore, the next steps and challenges of the project, particularly related to the identification of 'benefits' of RRI, will be discussed.

The MoRRI consortium is led by the Technopolis Group (Viola Peter, Nikos Maroulis), and further consists of: Center for Science and Technology Studies, University Leiden (Ingeborg Meijer); Aarhus University (Niels Mejlgaard); INGENIO CSIC-UPV (Richard Woolley, Ismael Rafols); Institute for Advance Studies Vienna (Erich Griessler, Angela Wroblewski); Fraunhofer ISI (Ralf Lindner, Susanne Buehrer); University College London (Jack Stilgoe); University of Athens (Lena Tsipouri).

References

Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Provencher, F., Rebout, L. and Roberge, G. (2014). Proportion of Open Access Papers Published in Peer-Reviewed Journals at the European and World Levels—1996–2013. Rapport, Commission Européenne DG Recherche & Innovation; RTD-B6-PP-2011-2: Study to Develop a Set of Indicators to Measure Open Access.

Barben, D.; Fisher, E.; Selin, C.; Guston, D. (2008): Anticipatory Governance of Nanotechnology: Foresight, Engagement, and Integration. In: Hackett, E.J.; Amsterdamska, O.; Lynch, M.; Wajcman, J. (Eds.): The Handbook of Science and Technology Studies, Cambridge, MIT Press. 979-1000

Bauer, M. W., Allum., N and Miller, S. (2007): What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public Understand. Sci.* 16, 79–95.

Björk, B-C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., and Guðnason, G. 2010. "Open Access to the Scientific Journal Literature: Situation 2009." PLoS ONE 5 (6): e11273. doi:10.1371/journal.pone.0011273.

Borrás, S. and Edler, J. (2014): Introduction: on governance, systems and change. In: Borrás, S. and Edler, J. (eds.): *The Governance of Socio-Technical Systems. Explaining Change*, Cheltenham: Edward Elgar Publishing, 1-22

Brom, F.W.A.; Chaturvedi, S., Ladikas, M., Zhang, W. (2015): Institutionalizing Ethical Debates in Science, Technology, and Innovation Policy: A Comparison of Europe, India and China. In: Ladikas, M.; S. Chaturvedi; Zhao, Y.; Stemerding, D. (Eds.) Science and Technology Governance and Ethics. Springer, Cham, Heidelberg New York, Dordrecht, London, 9-23. DOI: 10.1007/978-3-319-14693-5_12

Bucchi, M. (2008). Of deficits, deviations and dialogues. In M. Bucchi & B. Trench (Eds.), *Handbook of Public Communication of Science and Technology* (pp. 57–

76). Routledge.
Callon, M. (1999). The role of lay people in the production and dissemination of scientific knowledge. *Science, Technology & Society*, *4*, 81–94.
Caprile, M., Addis, E., Castano, C., Linge, I., Larios, M., Moulders, D. & Vazquez-Cuperio, S. ((2012): Meta-analysis of gender and science research, Synthesis report, Luxembourg, Publications Office of the European Union. Online: https://ec.europa.eu/research/science-society/document_library/pdf_06/metaanalysis-of-gender-and-science-research-synthesis-report.pdf
Catalyst (2004), The bottom Line: Connecting Corporate Performance and Gender Diversity, New York, San Jose, Toronto.
Costas, R., Meijer, I., Zahedi, Z., and Wouters, P. (2013). The value of research data - Metrics for datasets from a cultural and technical point of view. A Knowledge Exchange Report (pp. 1–48).
Davis, K.E., Fisher, A., Kingsbury, B., Merry, S.E. (2012). *Governance by Indicators. Global Power through Quantification and Rankings*. Oxford University Press, Oxford. DOI: 10.1093/acprof:oso/9780199658244.001.0001
Den Hertog, P., Jager, C.J., van Kan, A., Te Velde, R., Veldkamp, J., Aksnes, D.W., Sivertsen, G., van Leeuwen, T., van Wijk, E. (2014). Science, Technology & Innovation Indicators. Publication number: 2010.056-1435. This report is electronically available at www.sti2.nl
Dewey, J. (1934). The Supreme Intellectual Obligation. *Science Education*, *18*, 1–4.
European Commission (EC 2012). Towards better access to scientific information: Boosting the benefits of public investments in research. COM(2012) 401 final.
European Commission (EC 2014). Communication from the Commission. Towards a thriving data-driven economy. COM(2014) 442 final
European Commission (EC 2015); Indicators for promoting and monitoring Responsible Research and Innovation, report of the expert group on policy indicators for responsible research and innovation
http://ec.europa.eu/research/swafs/pdf/pub_rri/rri_indicators_final_version.pdf
European Union (2012): Responsible Research and Innovation. Europe's ability to respond to societal challenges. Available at: http://ec.europa.eu/research/sciencesociety/ document_library/pdf_06/responsible-research-and-innovation-leaflet_en.pdf
Farhan, H., Alonso, J., Davies, T., Tennison, J., Heath, T., & Berners-lee, T. (2013). Open Data Barometer, 1–45.
Felt, U., Fochler, M.; Müller, A., Strassnig, M. (2009): Unruly ethics: on the difficulties of a bottom-up approach to ethics in the field of genomics. Public Understanding of Science 18 (3), 354-371.
Fisher, E., et al., Mahajan, R.L., Mitcham, C. (2006). Midstream Modulation of Technology: Governance From Within. Bulletin of Science, Technology & Society Vol. 26, No. 6, December 2006, 485-496 DOI: 10.1177/0270467606295402
Friedman, B. (1996): Value-sensitive design. Interactions. Volume 3 Issue 6, Nov./Dec. 1996 pp. 16-23 doi:10.1145/242485.242493
Glerup, C. and Maja Horst, M. (2014) Mapping 'social responsibility' in science, Journal of Responsible Innovation, 1:1, 31-50, DOI: 10.1080/23299460.2014.882077
Grießler, E., Littig, B. (2006): Neosokratische Dialoge zu ethischen Fragen der Xenotransplantation. Ein Beitrag zur Bearbeitung ethischer Probleme in partizipativer Technikfolgenabschätzung. Buchinger, E.; Felt, U. (Hrsg.): Technik- und Wissenschaftssoziologie in Österreich. Stand und Perspektiven.ÖZS, Sonderheft 8/2006, 131-157.
Grin, J., Grunwald, A. (2000, eds.): Vision Assessment: Shaping Technology in 21st Century Society. Heidelberg: Springer 2000, 193 S.
Guston, D.H. and Sarewitz, D. (2002) Real-time technology assessment. Technology in Society 24(1–2): 93–109.
Guston, D. H. (2014): Understanding 'anticipatory governance'. In: Social Studies of Science, 4482:218-242
Hagendijk, R., & Irwin, A. (2006). Public deliberation and governance: engaging

with science and technology in contemporary Europe. Minerva, 44(2), 167-184.

Hellström, T. (2003): Systemic innovation and risk: technology assessment and the challenge of responsible innovation. In: *Technology in Society* 25 (3), p. 369–384

Karinen R. and Guston, D.H. (2010) Toward anticipatory governance: The experience with nanotechnology. In: Kaiser M, Kurath M, Maasen S, Rehmann-Sutter C (eds)*Governing Future Technologies: Nanotechnology and the Rise of an Assessment Regime*. Dordrecht: Springer, pp. 217–232.

Kuhlmann, S., Edler, J., Ordóñez-Matamoros, G., Randles, S., Walhout, B., Gough, C., Lindner, R. (2016): Responsibility Navigator. In: Lindner, R.; Kuhlmann, S.; Randles, S.; Bedsted, B.; Gorgoni, G.; Griessler, E.; Loconto, A.; Mejlgaard, N. (eds): Navigating Towards Shared Responsibility in Research and Innovation. Approach, Process and Results of the Res-AGorA Project, Karlsruhe, p. 135-158 https://indd.adobe.com/view/eaeb695e-a212-4a34-aeba-b3d8a7a58acc

Ladikas, M.; S. Chaturvedi; Zhao, Y.; Stemerding,D. (Eds. 2015) Science and Technology Governance and Ethics. Springer, Cham, Heidelberg New York, Dordrecht, London.

Lindner, R.; Warnke,P.; Kukk, P. (2015): Update of the literature review & Visioning exercise. Progress Report D5.1, Deliverable of the Project Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI), Brussels.

Lipinsky, A. (2014), Gender Equality Policies in Public Research, Luxembourg, Publications Office of the European Union. Online: http://ec.europa.eu/research/pdf/199627_2014%202971_rtd_report.pdf

McKinsey and Company (2007). Women Matter: Gender diversity, a corporate performance driver, McKinsey & Company Inc. Online: http://www.mckinsey.de/sites/mck_files/files/Women_Matter_1_brochure.pdf

Mejlgaard, N. and Griessler, E. (2016): Monitoring RRI in Europe: approach and key observations. In: Lindner, R.; Kuhlmann, S.; Randles, S.; Bedsted, B.; Gorgoni, G.; Griessler, E.; Loconto, A.; Mejlgaard, N. (eds): Navigating Towards Shared Responsibility in Research and Innovation. Approach, Process and Results of the Res-AGorA Project, Karlsruhe, p. 115-118. https://indd.adobe.com/view/eaeb695e-a212-4a34-aeba-b3d8a7a58acc

MoRRI (2015). A series of documents produced as part of the European Commission's service contract RTD-B6-PP-00964-2013, "Monitoring the evolution and benefits of responsible research and innovation". The following deliverables have been produced so far:

- Analytical Report on the Dimension of Citizen Engagement and Participation of Societal Actors in Research and Innovation. Sub-task 2.5, deliverable D.2.1.
- Analytical Report on the Gender Equality Dimension. Sub-task 2.5, deliverable D.2.3.
- Analytical Report on the Dimension of Science Literacy and Scientific Education. Sub-task 2.5, deliverable D.2.2.
- Analytical Report on the Dimension of Open Access. Sub-task 2.5, deliverable D.2.4.
- Analytical Report on the Dimension of Research and Innovation Ethics. Sub-task 2.5, deliverable D.2.4.1.
- Analytical Report on the Dimensions of Research and Innovation Governance. Sub-task 2.5, deliverable D2.4.2
- Synthesis report on existing indicators across RRI dimensions, Task 3, Progress report, Deliverable D3.1, May 2015
- Metrics and indicators of Responsible Research and Innovation, Task 3, Progress report, Deliverable D3.2, September 2015

Nielsen, M. V.; Lindner, R.; Bryndum, N.; Burchardt, U.; Schofield, M.; Stilgoe, J. (2015): Navigating towards Responsible Research and Innovation. Challenges for Policy and Governance. In: Michalek, T.; Scherz, C.; Hennen, L.; Hebáková, L.; Hahn, J.; Seitz, S. (Hg.): The Next Horizon of Technology Assessment. Proceedings from the PACITA 2015 Conference in Berlin, Berlin, S. 57-62

Owen, R., Macnaghten P., Stilgoe J. (2012) Responsible research and innovation: From science in society to science for society, with society. Science and Public Policy 39; 751–760.

Pandza, K. and Ellwood, P. (2013). Strategic and ethical foundations for responsible innovation. Res. Policy 42(5); 1112–1125 doi:10.1016/j.respol.2013.02.007

Polanyi, M. (2000). The republic of science: Its political and economic theory. Minerva, 38(1), 1-21.

Randles, S.; Laredo, P.; Loconto, A.; Walhout, B.; Lindner, R. (2016): Framings and frameworks: Six grand narratives of *de facto* rri. In: Lindner, R.; Kuhlmann, S.; Randles, S.; Bedsted, B.; Gorgoni, G.; Griessler, E.; Loconto, A.; Mejlgaard, N. (eds): Navigating Towards Shared Responsibility in Research and Innovation. Approach, Process and Results of the Res-AGorA Project, Karlsruhe, p. 31-36. https://indd.adobe.com/view/eaeb695e-a212-4a34-aeba-b3d8a7a58acc

Ravn, T., Mejlgaard, N. and Rask, A. (2014): Inventory of PE mechanisms and initiatives D.1.1. Available at: http://pe2020.eu/wp-content/uploads/sites/15/2014/02/PE2020-FINALD.1.1-report.pdf

Rip, A. (2014): The past and future of RRI. Life Sciences, Society and Policy 10:17 http://www.lsspjournal.com/content/10/1/17

Rip, A., Misa, T.J., Schot, J.W. (1995). *Managing technology in Society. The approach of constructive technology assessment*. Pinter Publishers, London.

Rowe, G. and Frewer, L.J. (2005): A Typology of Public Engagement Mechanisms. *Science Technology & Human Values*. 30: 251.

Schicktanz, S.; Schweda, M., Wynne, B. (2012): The ethics of ‚public understanding of ethics' – why and how bioethics expertise should include public and patients' voices. Med Health Care and Philos 15: 129-139.

Schiebinger, L. and Schraudner, M. (2011), Interdisciplinary Approaches to Achieving Gendered Innovations in Science, Medicine, and Engineering, Interdisciplinary Science Review, Vol. 36, No. 2, 154–67.

Shapin, S. (2008): The Scientific Life. A moral history of late modern vocation. The University of Chicago Press.

Stilgoe, J., Owen, R., Macnaghten, P. (2013) Developing a framework for responsible innovation. Res. Policy 42;1568-1580, http://dx.doi.org/10.1016/j.respol.2013.05.008

The Royal Society. (1985). *The Public Understanding of Science*. Retrieved from https://royalsociety.org/~/media/Royal_Society_Content/policy/publications/1985/10700.pdf

Van den Eynden, V. and Bishop, L. (2014). Sowing the Seed: Incentives and motivations for sharing research data, a researcher's perspective. A Knowledge Exchange Report.

Von Schomberg, R. (2013). "A vision of responsible innovation". In: R. Owen, M.Heintz and J Bessant (eds.) *Responsible Innovation.* London: John Wiley.

Wickson, F., and Carew, A. (2014). Quality criteria and indicators for responsible research & innovation. Learning from transdisciplinarity. Journal of Responsible Innovation, DOI: 10.1080/23299460.2014.963004

Wilsdon, J. and Willis, R. (2004) *See-through science: why public engagement needs to move upstream.*Project Report. Demos, London.

Wilsdon, J., et al. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. DOI: 10.13140/RG.2.1.4929.1363

Habermas, J. (1987). The theory of communicative action. Volume II Lifeworld and system: a critique of functionalist reason. Boston Beacon Press.

Wynne, B. (1995). Public Understanding of Science. In Jasanoff, S. *et al*. (eds.). *Handbook of Science and Technology Studies*. Sage Publications, London.