Grant Agreement Number: 824671

SUPER MoRRI – Scientific understanding and provision of an enhanced and robust monitoring system for RRI

D1.2 A Monitoring Framework for Responsible Research and Innovation
Strategic Development Plan 2020-24

Authors: Richard Woolley, Paula Otero-Hermida, Niels Mejlgaard, Thomas Ryan, Kjetil Rommetveit, Roger Strand, Wouter van de Klippe.

Submission date: 14.07.2020

Version: 2

Type: Report

Dissemination Level: Public

Project website: www.supermorri.eu

A Monitoring Framework for Responsible Research and Innovation

Strategic Development Plan 2020-24

EXECUTIVE SUMMARY

The SUPER MoRRI project is guided by three strategic documents, including this Strategic Plan, the Implementation Plan (D2.1) and the Case Research Plan (D5.1). The Strategic Plan makes clear the understanding of the task and the principles guiding the development of the SUPER MoRRI approach.

Responsible Research and Innovation (RRI) is an approach to institutional change that emphasises the vital importance of engagement between actors and stakeholders from across the spectrum of science, innovation and society. The task of SUPER MoRRI is to support processes of institutional change through the provision of a monitoring framework that can promote learning about transformation processes relevant to diverse stakeholders (including policymakers), build and reinforce collaboration between all stakeholders including citizens, and allow for appropriate and credible comparisons at a variety of levels of analysis.

SUPER MoRRI is concerned that the data and information gathered and presented through the monitoring framework should be useful for different stakeholders and potential users. These users should therefore be consulted in the processes of designing, developing and presenting data and information, including indicators.

SUPER MoRRI will seek to innovate in its approach to responsible quantification, by presenting data and information with supporting tools that promote the making of appropriate comparisons and interpretations by users. SUPER MoRRI intends to develop an approach to the 'credible contextualisation' of data and information, including indicators, such that a clear understanding of the degree of context dependence of the information provided is available to support users' interpretations.

A critical inspection was undertaken of the indicators developed in the MoRRI project to monitor the evolution and benefits of RRI at the Member State level. These indicators were assessed in terms of four criteria: relevance; validity; feasibility; and usefulness. A number of MoRRI indicators will be adopted unchanged in SUPER MoRRI, while others will be modified so as to satisfy SUPER MoRRI criteria. A project Working Paper detailing the critical inspection of the MoRRI indicators is attached as an Appendix to this document.

SUPER MoRRI plans to make use of, or further develop, existing 'indicators in the wild' from sources such as She Figures, Eurobarometer and the Open Science Monitor, for example. Ongoing scanning of the research and policy environment to ensure awareness of potentially useful data or indicators will continue throughout the project.

The development of new data and information for monitoring RRI will include the utilisation of a Country Correspondents Network (CCN) in Europe. The CCN will participate in research activities involving research funding organisations and research performing organisations.

A group of International Satellite Partners (ISP) have been included in the project and will ensure learning from the global dimension of RRI. ISPs are based in Argentina, Australia, Brasil, Canada, PR China, Iran, Japan, South Africa and the USA.

An innovative 'SwafS project eco-system' will operate across the life of the SUPER MoRRI project to co-create data and information that supports practitioners' needs. This will include a form of self-assessment tool that can help participants monitor and reflect on implementation processes and the challenges of longitudinal development of RRI initiatives.

SUPER MoRRI will implement an original research programme of studies that will produce new indicators of patterns of RRI activities and of the outcomes that mark pathways toward benefits from RRI. Details of the SUPER MoRRI research programme can be found in the Implementation Plan (D2.1) and the Case Research Plan (D5.1).

The SUPER MoRRI project consortium recognises the need for reflexivity about its own assumptions, not least due to uncertainty regarding future policy settings and evolving framings of desirable approaches to institutional change. A broad vision of responsibility in research and innovation has been adopted by SUPER MoRRI to maximise the potential of the project outputs and outcomes to have value beyond the lifespan of the project and of the SwafS Work Programme.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Governance in *complexity, not governance* of *complexity*

The aim of this document is to contribute to the development of a monitoring framework for responsible research and innovation (RRI). This strategic plan sets out the broad principles and some areas of activity of the SUPER MoRRI research project for the period 2020-24. It is a companion piece to the Implementation Plan 2020-24, which sets out the planned data vehicle collections for SUPER MoRRI, and the Case Research Plan 2020-24 that details the planned case studies. These plans will be updated periodically throughout the life of the SUPER MoRRI project according to milestone events, a process of mid-term reflection and review, and new developments emerging from the work of the project and beyond. The three planning documents provide conceptual and research programme pillars that will structure the definition, selection and operationalisation of our research questions.

Figure 1: SUPER MoRRI project development: three pillars
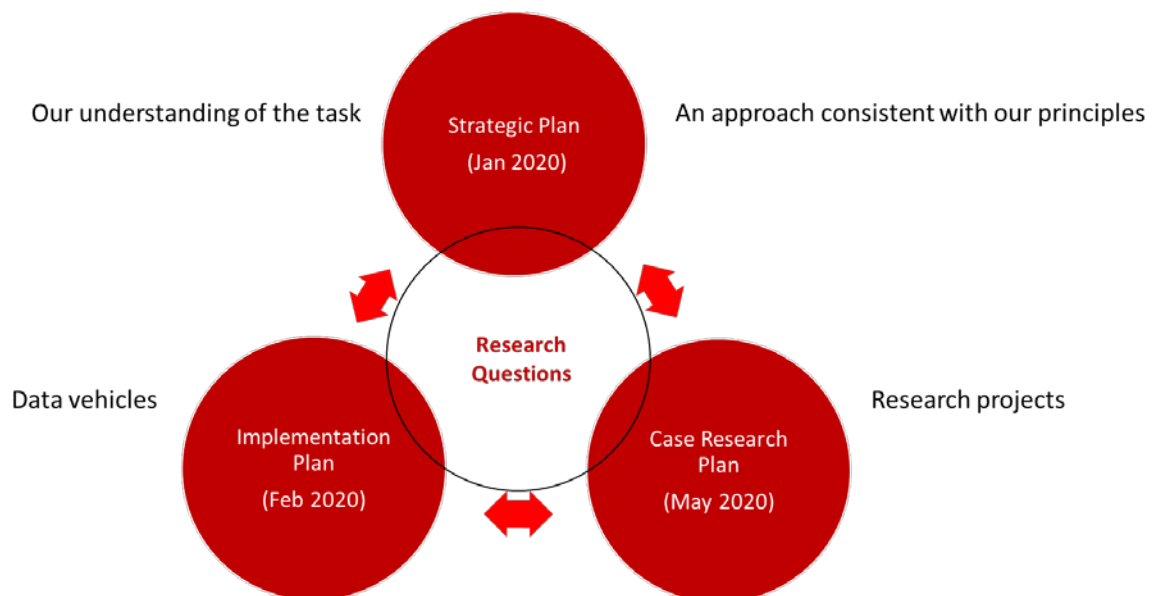
Principal elements in the development of the strategic plan have been a review of existing approaches and tools for monitoring RRI, a stakeholder workshop, expert consultations, engagement with current SwafS projects and the circulation of a call for comments on a prior Briefing Paper. In the sphere of academic research, the plan builds on key conceptual and empirical contributions to develop a coherent overall approach to RRI (section 2). In the sphere of policy research, the plan builds on previous work conducted by the Expert Group on Policy Indicators for Responsible Research and Innovation (Strand et al. 2015), the Expert Group on Altmetrics (Wilsdon et al. 2017), the Open Science Policy Platform (Hormia-Poutanen et al. 2017), the MoRRI project (Mejlgaard et al. 2018; Peter et al. 2018), She Figures (2019) and the Expert Group on Indicators for Researchers' Engagement with Open Science (Wouters et al. 2019). This initial version of the strategic plan is viewed as a working document that will likely undergo considerable modification, including milestone updates, continued conceptual development and the integration of emerging activities, leading up to the delivery of a monitoring framework for RRI and associated user tools by 2024.

A key factor in the development of this strategic plan is the apparent strong support among stakeholders for more responsible research and innovation, including stakeholders within the research and innovation (R&I) system, users of the knowledge and other outputs of the R&I system and citizens in general. A monitoring framework can be seen as one element that can support and influence transformations of R&I in the interests of placing a higher value on responsibility. Of course, the degree of complexity that exists within the organisation of science and research means that the general support for responsible practices and outcomes in knowledge production can have quite specific characteristics in different contexts, such as diverse epistemic communities. The same complexity characterises different technology fields or 'wicked problems'. This strategic plan therefore sets out to support RRI by identifying relevant levels of analysis for gathering data and coupling this with a clear

understanding of appropriate forms and degrees of de-contextualisation for interpretations of these data. This process will involve experimentation in data gathering and the design of interpretive models that will identify opportunities for monitoring. Wherever the development of quantitative indicators is under consideration we propose to engage in indicator co-creation, involving groups of users of quantification tools within the monitoring framework. In this way we will try to ensure the overall goal of supporting RRI is not hindered or negatively affected by inappropriate quantification innovations, or unnecessarily complicated by the creation of 'indicators for indicators sake' that do not serve users' monitoring purposes. In this sense the outline of a monitoring framework for RRI contained in this planning document considers taking a responsible approach to the design and uses of indicators (or other quantification tools) as a paramount value (Hicks et al. 2015; Wilsdon et al. 2015).

Conceptual debate regarding RRI is ongoing, energising and desirable. From our perspective the diverse existing theoretical approaches to RRI represent not a conceptual conflict in need of resolution but a range of ideas that generate important principles and identify substantive issues that can contribute to our design approach. The fundamental point, as Fisher (2018: 53) describes, is that "deep structural ambivalence to science and expertise mean that responsible innovation as a necessary aspiration is here to stay". Our basic motivations to develop a monitoring framework for RRI follow from this recognition. First, more responsible knowledge production and innovation regimes are both desirable and necessary, and second, a reflexive process of information gathering, sharing and learning can support and contribute to this aspiration.

2. THE CONCEPTUAL HETEROGENEITY OF RESPONSIBLE RESEARCH AND INNOVATION

Conceptual thinking regarding RRI embraces a variety of approaches, which have developed more or less in parallel rather than through contestation over a hegemonic definition or understanding (European Commission 2014, 2017; Owen et al. 2012; Stilgoe et al. 2013; Von Schomberg 2013, 2014). A recent review of the underlying ontological and axiological assumptions of the main conceptualisations of responsible innovation (Timmermans and Blok 2018) argued that differences among these visions are in part tied to the policymaking and scholarly contexts from where they emerged. This can be viewed as a strength in terms of potential versatility of operationalisations available in the conceptual toolbox of RRI, but also as a weakness in terms of fragmentation of researcher and practitioner communities and the lack of coherence in measures to realise the aspiration for greater responsibility in both scientific knowledge production and innovation (Timmermans and Blok 2018).

2.1 Three Conceptualisations of RRI

The European Commission (EC) conception of RRI emerged from its science with and for society (SwafS) policies and R&D work programmes. The EC describes RRI as diverse sets of societal actors (researchers, citizens, policymakers, business, third sector organisations, etc.) that "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 society".[1] The major mechanism for bringing actors together is public engagement, one of the EC's six RRI 'keys' along with ethics, gender equality, governance, open science and science education. Implementing interventions focused on each of these keys thus describes the broad pathway toward better alignment between science, innovation and society. The abstract framing of the EC definition around 'alignment' is to some extent counterbalanced by the six keys as categories that can be more straightforwardly operationalised. Nevertheless, how, and how much, actions on the various keys translate in terms of alignment is not clear. Although there has been no official policy move to extend the number or scope of RRI keys, the EC Expert Group on Policy Indicators for Responsible Research and Innovation (Strand et al. 2015) recommended the addition of thematic keys for 'social justice/inclusion' and 'sustainability', which perhaps provides one clue in this direction.

Rene Von Schomberg's well-known conception of RRI describes "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 and its marketable products (in order to allow a proper embedding of scientific and technological advances in

---

[1] https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation

our society" (Von Schomberg 2014: 63). The need for collective responsibility and 'mutual responsiveness' is driven by a series of global deficits in research and innovation (Von Schomberg 2019). In particular, the reproducibility and (over)production crises in science, innovation governance focused on macro-economic impact, regulation concerned exclusively with risks, and markets obsessed with technological potential, form a self-reinforcing system that is underdetermined in its orientation toward broadly shared public values and expectations. The normative turn towards responsibility is thus a consequence of the "directionality failure" of innovation systems and innovation policy (Lindner et al 2016). Von Schomberg (2019) argues strongly that the "right impacts" from R&I are already obvious, linked to strategies for addressing 'grand societal challenges' and embedded in multi-party agreements such as the Lund Declaration 2009. Overcoming the multidimensional failure of research and innovation requires a collective response that is transparent about where scientific and societal certainty is lacking and prepared to 'open up' innovation pathways to multiple alternative outcomes and solutions (Stirling 2007), understanding individual or sectoral interests as nevertheless mutually constituted and interdependent.

The third prominent conceptualisation of RRI, described by Jack Stilgoe, Richard Owen and Phil Macnaghten, defines RRI as "taking care of the future through collective stewardship of science and innovation in the present" (Stilgoe et al. 2013: 1530). The authors operationalise collective stewardship through four 'dimensions' anticipation, reflection, inclusion and responsiveness. Individual and collective practices that institutionalise these dimensions can be expected to reflexively shape the purposes, processes and products of scientific research and technological innovation. Collective stewardship should also seek to include, within the innovation process, a range of knowledges that are wider than engineering and technical knowledge, such as indigenous, ethical and social science knowledge. Whereas the European Commission and Von Schomberg both anticipate benefits from interactions among diverse actors throughout the R&I process, Stilgoe and colleagues consciously reject a "consequentialist framing for responsibility" (Stilgoe et al. 2013: 1569) in favour of an ethics focused on the qualities of processes (care, custodianship).

Despite their differences, the three most prominent conceptualisations of RRI share a number of core precepts:
● the inclusion of both scientific research and technological innovation within the framing of what needs to be transformed;
● a fundamental reliance on the integration of diverse scientific, innovation and societal actors as the transformative (social) mechanism;
● the imperative to always look beyond scientific and technological challenges, problems and opportunities to concurrently consider societal uncertainties;
● the value of diverse domains and types of knowledge; and
● a requirement that research and innovation practices and processes be more transparent, accessible and inclusive.

We summarise these prominent conceptualisations of RRI within a three-part model of integration, implementation and impact (i3).

- **Integration** of diverse actors, knowledges, capabilities and interests, refers to the mobilisation of individuals, organisations, institutions, technology and resources for R&I. Integration occurs at multiple levels of organisation and with varying scope. It includes formal vehicles such as strategic alliances, contracts, and projects and informal arrangements of cooperation. Integration is relatively responsible when it is plural, diverse and inclusive.
- **Implementation** refers to collective research and innovation processes and practices. Implementation pathways are relatively responsible when based on negotiated and interdependent goals, mutual commitment to avoiding adverse social, environmental and other effects, and shared (normative) expectations regarding users and beneficiaries.
- **Impact** therefore refers principally to transformations in processes, connections, capacities, attitudes, identities and anticipated possible futures, rather than to the outputs and outcomes of R&I they carry.

The model does not represent a sequenced or linear approach to either understanding or monitoring responsible research and innovation. The three categories in the model are entangled with each other. Implementation processes can lead to the integration of new actors or knowledge; impacts generate feedback processes that can modify implementation or reconfigure integration and so on.

From a monitoring perspective the three analytical categories in the i3 model are designed to help frame questions for monitoring to support RRI. Each category prompts different types of questions, for example:

- Integration: who is involved? on what basis is the participation of diverse actors organised? what types of knowledge and technology are involved? how are citizens involved?
- Implementation: how are interactions organised? how are priorities developed and agreed? how are conflicts between goals, or among priorities, exposed, debated and resolved? Are processes transparent and activities open and inclusive? how are emerging scientific controversies, technical obstacles and/or societal uncertainties treated? are multiple innovation pathways generated and developed?
- Impact: how do users of research and innovation provide feedback to knowledge producers and innovators? are potential beneficiaries included in the R&I cycle and at what point? how do networks of users transmit and modify innovations? do beneficiaries have the potential to become users? where do these translations spread and who do these networks include? how are emergent effects of innovations governed by users and beneficiaries and communicated to producers, innovators and/or regulators?

These questions become more meaningful when it is assumed that research and innovation is always-already oriented toward a **problem area**. Research and innovation never starts from a blank slate but is continuously oriented by visions, questions and challenges, by prior scientific findings and successful and unsuccessful innovations. Rommetveit and colleagues (2019) also frame responsible innovation in terms of networks of actors and stakeholders evolving in contexts of uncertainty and argue that the focus of these networks is (and should be) on problems. The configuration of RRI in any empirical context will thus necessarily reflect understandings of how to address the epistemic, technical and societal uncertainties involved in transforming problem areas of varying scales and scope (Callon et al. 2009; Funtowicz and Ravetz 1993).

Problem articulations may take place in several places: as policy agenda setting; in research teams trying to address specific problems, or by the users of a given product or process (see section 3.2). In the case of indicators, the user perspective is of major importance, as the question of what problem is being addressed and assessed through indicators, must necessarily be included in the process. Problem-orientation of such activities can serve to specify collective purposes (see section 3.1), to coordinate and orchestrate contributions from differently positioned scientific fields or disciplines, and to integrate data collected from various sites and sources. This means that problem articulations must be constantly revised and re-assessed throughout all (or most) stages of research, innovation and policy processes, and so should be considered in all integration, implementation and impact processes.

Table 1: Conceptual summary guiding monitoring framework development

	Responsible Research and Innovation		
	Integration	*Implementation*	*Impact*
European Commission	Diverse societal actors	Public engagement in the whole research and innovation process	Increased alignment between R&I and society
Von Schomberg	Societal actors and innovators	Mutual responsiveness	Better marketable products Liberal-democratic values
Stilgoe, Owen & Mcnaghten	Collective stewardship	Anticipation, reflection, inclusion, responsiveness	

2.2   Narratives of Transformation ⇆ Theories of Change

The major conceptual characterisations of RRI, described above, provide broad expectations of the types of transformations that working toward more responsible R&I could, or should be concerned to, deliver. Strand and Rommetveit (2019: 4-7) summarise the narrative structures and transformative ontologies that have been constructed around RRI. They identify four narratives:

A. **Regaining control over the runaway train of science and technolog**y, which is a two-edged sword that provides us with benefits, goods and welfare, but also with Hiroshima, Silent Spring, Chernobyl, designer babies, enhanced soldiers, CRISPR, killer drones, Facebook, Cambridge Analytica, etc. In the narrative of the runaway train, RRI is conceived as (part of) the solution by 1) **acting as handbrake** through external regulations such as ethical reviews, and through soft self-governance such as researchers' anticipation of risks and benefits of their work; or 2) **switching or re-directing the train** through processes such as co-creation, democratization and upstream engagement in research agenda setting.

B. **Research and innovation is the locomotive force** of the (capitalist) knowledge economy and that the **train is on the tracks** towards economic growth, increased human welfare and progress. One variant of this narrative is that the train is on the tracks but ungrateful **citizens are obstacles in its way, raising barriers and protesting in the middle of the railroad** (active version) or hindering and slowing it down through psychological inertia, ignorance and lack of knowledge (passive version). **From this perspective, RRI is a solution to the questions: How do we educate, reassure and calm down the ignorant public and make them trust us, trust science again?** And in this perspective, the "5 keys" make perfect sense.

C. The official policy narrative, that motivated RRI as a cross-cutting principle of Horizon 2020, implements **RRI to improve the alignment of science with society**. RRI can be seen as part of a much broader development that aims at **improved communication and collaboration** across epistemic communities (such as disciplines and research fields) but also across entire sectors such as research, technology development, politics and law… are **also attempts at creating legitimacy**, such as when ethics, law, technology assessment or public engagement expertise is coupled onto biotechnology, nanotechnology, or ICT projects. The central metaphor of this narrative is not so much one of R&I as a linear train as one of nonlinear **networks of actors, institutions and expertise.**

D. Rather than R&I being the locomotive and RRI the momentum for change, **RRI itself is the train.** While this narrative has little power in general, projects like SUPER MoRRI and academics (such as us) consolidate and reify RRI as "something" through our own activities, by writing RRI books, creating journals (such as the *Journal of Responsible Innovation*), teaching RRI courses and so on. In this way we are creating our own epistemic network around RRI. The fact that

SUPER MoRRI is itself a (publicly funded) stakeholder should be continually recalled and reflexively factored into the approach to monitoring RRI we develop.

These narratives provide long interpretive arcs that help us understand how the institutionalisation of RRI as a policy approach is imagined to transform science, innovation and society. They provide powerful thinking tools in the context of working toward a monitoring framework for RRI by shaping our understanding of transformation dynamics at a discursive level, particularly as this relates to explicit 'policy-driven RRI' initiatives. They are also important for addressing the above-mentioned problem-orientation of RRI, since each narrative is likely to be formed around different problem understandings and articulations.

Concurrently, at the level of policies, programmes, projects and other actions, a relatively meso or micro dynamic of transformation characterises emergent thinking and acting designed to transform R&I. Scientist- and citizen-driven movements for open science, for example, are motivated to reconfigure the production, accessibility and usability of knowledge. Such initiatives are based on an understanding that change is needed. They assume that certain actions are required to bring these necessary changes to reality. The SUPER MoRRI monitoring framework is also designed to be inclusive of such activities, which whilst they may not be linked to the institutionalisation of policy-driven RRI at all, are intimately engaged in the 'responsibilisation' of research and innovation and change in the relationship between science and society - often at the frontier of transformation processes.

Whether explicitly linked to RRI policy and practice, or emerging organically, such initiatives are motivated by a desired transformation and the formal or informal formulation of logics and pathways of actions to achieve it, and guided by an understanding (explicit and/or implicit) of the policy or innovation problem at hand. For example, proponents of closer relationships between science and society might argue for requiring a citizen science component in every ERC Starting Grant. Advocates for more relevant outputs from R&I might lobby for collectively defining potential innovation pathways even prior to a scientific breakthrough being made. The multiplicity of theories of change at the level of practice reflect varying understandings of what transformations might be expected and how these are linked to the specifics of our thought and action. The interactions between relatively macro-level discursive formations that conceptualise systemic transformation processes and relatively micro-level theories of change at the level of practices are mutually entangled and constitutive of each other.

Within the SUPER MoRRI monitoring framework, different data elements can be expected to reflect interventions or studies guided by a variety of theories of change. These theories of change may be articulated either explicitly or implicitly. Tracing and revealing how a theory of change institutionalises a new set of arrangements as planned, or deviates, loop backwards, creates perverse incentives, and so on, can be a critical piece of learning for users of monitor. SUPER MoRRI will seek to provide relevant guidance, wherever this is possible and feasible, regarding valid and reasonable theories of change

linked to specific contexts of data gathering. The aim here is to support users' capacities to make informed interpretations of the data and information provided by SUPER MoRRI.

3. A MONITORING FRAMEWORK FOR RRI

SUPER MoRRI is designed as a **monitoring framework** in which a variety of different forms of information gathering at different levels of analysis co-exist, with the aim of supporting transformation in R&I toward enhanced responsibility, for the benefit of R&I and society.

3.1 Purposes and functions

Processes of continuous observation of social phenomena, including highly complex entities such as research and innovation, are multi-faceted. Monitoring alone can have multiple purposes including:

- learning about transformation processes and informing ongoing policy design and response;
- building and reinforcing collaboration and trust among policymakers, stakeholders and citizens;
- accountability of policymakers, programme and project managers (Kleibrink et al. 2016);
- making appropriate comparisons, including bench-marking; and
- reflexively engaging with our own assumptions.

SUPER MoRRI is being designed as a **monitoring framework** that can be useful for a number of different purposes. We prefer a 'framework' as a model that is open to accommodating an array of different elements that contribute to these monitoring purposes. A monitoring framework allows for *bricolage*, experimentation, discrete perspectives and admits potentially high degrees of context dependence. In contrast, we understand a monitoring system as driving toward standardisation, a totalising vision and context independence.

Monitoring has three main functions:

- gathering valid information that can be considered by decision-makers at all levels;
- making clear the aims and functioning of policy and development strategies for all stakeholders and the public; and
- ensuring transparency that supports the involvement and participation of stakeholders (Kleibrink et al. 2016).

SUPER MoRRI will gather data and information at different geo-spatial and institutional levels that are relevant to decision-makers and to stakeholders and citizens more broadly. For example, within the research system it will gather data and information at scientific field, discipline, research group and individual levels. It will incorporate elements that support responsibility in R&I from a variety of stakeholder perspectives, providing resources that foster understanding, learning and inclusivity. It is for this reason that we have underlined problem definition and purpose specification as centrally important activities, since the monitoring needs of users will vary (with time, place, context, etc.) and cannot be determined once and for all.

3.2   Users

Monitoring only has value in use. Different users of the framework will likely have varying principal or secondary purposes for monitoring. The same data and information elements may have different value and meaning depending on users' purposes. SUPER MoRRI intends to develop guidelines for using and interpreting the quantifications presented, including indicators. Relevant contextual information will also be a necessary complement to these tools.

A number of key users for a monitoring framework for RRI can be anticipated, although these should not be considered an exhaustive set.

- Science, research and innovation policymakers - including national and regional agencies
- Higher education policymakers - including national and regional agencies
- Research managers, directors and planners at research performing organisations (RPOs) - including universities and national research institutes
- RRI practitioners - including in academia and regional development
- Research funding organisations - including public agencies and private foundations
- Science with and for Society actors - including citizen science organisations
- Companies - including R&D performing firms and technology and social innovation focused SMEs
- Third sector organisations working in knowledge-intensive fields or deploying scientific or technological know-how
- Researcher development related organisations - including professional development agencies and firms, accreditation agencies, research evaluators, human resource units in RPOs

The data and information presented the monitoring framework will be developed in collaboration with users. This includes a co-creation phase wherever indicators, visualisations or other vehicles for communicating monitoring are being prepared (section 4.5).

3.3   Responsible Quantification

Support for the capacity to make well-informed interpretations of data and information reflects the SUPER MoRRI commitment to responsible quantification. This indicates that quantification and monitoring are not mere means towards the ends of providing 'Responsibility', but that quantification itself should be undertaken in accordance with an RRI approach. A variety of quantification techniques will be used to provide users with credible and useful monitoring outputs within the framework. These techniques will include data visualisation coded to European geo-spatial regions such as NUTS2/3, science maps of research outputs, network diagrams, indicators and a range of charts, figures and graphs. Primary data underpinning tools and resources provided in SUPER MoRRI will be available to

prospective users under FAIR (findable, accessible, interoperable, reusable) principles. Guidance on the responsible use of these data will also be provided.

SUPER MoRRI will pioneer a responsible quantification approach called **credible contextualisation**. There are no universal context-free indicators or other data quantifications of RRI or related concepts. Rather, data used in indicators are gathered in a specific context. The degree to which any quantification can be utilised as a comparator or as a benchmark, for example, depends on the degree of de-contextualisation this quantification can credibly stand. All quantifications in SUPER MoRRI will therefore be accompanied by a **credible contextualisation band or scope recommendation** to support users to understand the contours of appropriate comparisons and limits for scalar (dis)aggregation, for example.

Specifically in relation to indicators, all metrics and measures on which indicators produced or used within the monitoring framework will be transparent. This approach will ensure that all measures used as proxies for a concept or category are known, and the extent to which their validity has been evaluated will be clearly described. This responsible quantification approach will aim to ensure that no monitoring elements rest on ecological fallacies.

3.4   Indicators

Indicators are quantification tools which 'indicate' in the sense of providing information about an entity which is not directly measurable and 'signal' in the sense that they must be interpreted to have meaning (Lehtonen 2017). What separates indicators from other data or information is that they are the product of a conceptual framework or model that justifies the choice of data and logic of their interpretation (theory of change). Indicators are a particular type of quantification, that due to their ready association with specific targets, goals or problems have become ubiquitous in communicating about policy, corporate performance, and condensed trendlines regarding complex socio-economic phenomena. Indicators are indeed ubiquitous.

Indicators are viewed as one information element within a monitoring framework for RRI. The SUPER MoRRI empirical research programme (section 4) will generate primary data that may in some circumstances be processed to construct one or more indicators. There are also a number of relevant sources for existing indicators that may contribute to the monitoring framework.

3.4.1   Indicators in the Wild

A significant number of indicators exist that appear relevant for a monitoring framework to support RRI. Some of these are 'indicators in the wild' (Ràfols 2018) that provide an important avenue for learning, as not only the indicator design can be considered but also the desirability of an indicator's

apparent effects. There is no need for SUPER MoRRI to reinvent the wheel where good quality indicators have already been designed and, in some cases, are also backed by an effective data gathering procedure. This section contains a brief summary of relevant indicators and indicator frameworks that are already in the wild.

MoRRI indicators of RRI

The MoRRI project produced a set of indicators based on the European Commission's keys (Peter et al. 2018). These indicators were based on a number of original data collections and cherry-picking relevant indicators in the wild from other sources. A core criteria guiding the development of the MoRRI indicators was that they should be aggregated at European Union Member State (MS) level and cover all MS. A critical inspection of the MoRRI indicators according to a number of assessment criteria was undertaken in SUPER MoRRI (see Appendix A). Table 2 summarises our provisional assessment of the MoRRI indicators considered relevant, valid, feasible and useful for the monitoring framework.

The MoRRI indicators were designed with a policy accountability purpose in mind. Several of the indicators developed appear potentially relevant for the potentially diverse monitoring purposes of SUPER MoRRI. A total of six indicators were considered to satisfy all of our critical assessment criteria (Table 2 column 1). Five of these indicators are for the gender equality key, and one is for the open access key.

A total of nineteen MoRRI indicators were assessed positively for some criteria but not others (Table 2 Column 2). Five of these indicators were for the gender equality key, two for the science literacy/education key, six for the public engagement key, and two each for the ethics, open access and governance keys. We consider each of the indicators to have the potential to add value to a monitoring framework to support RRI. However, modifications to these indicators would be required, including changes to methods and levels of analysis for data collection and the specification of the metric.

Table 2: MoRRI indicators in a monitoring framework for RRI

Relevant, valid, feasible and useful	Possibly, with modifications
Share of female researchers, all sectors	Share of research-performing organisations with gender equality plans
Dissimilarity index, higher education	Share of research-funding organisations (RFOs) promoting gender content in research
Glass ceiling index	Share of research-performing organisations (RPOs) with policies to promote gender in research content
Gender wage gap	Share of female heads of research-performing organisations
Share of female authors	Share of gender-balanced recruitment committees at research-performing organisations
Share of open access publications	RRI-related training at higher education institutions
	Citizen science activities in research-performing organisations
	Policy-oriented engagement with science
	Citizen preferences for active participation in science and technology decision-making
	Active information search about controversial technologies
	Public engagement performance mechanisms at the level of research-performing organisations
	Embedment of public engagement activities in the funding structure of key public research-funding agencies
	Public engagement elements as evaluative criteria in research proposal evaluations
	Ethics at the level of research-performing organisations
	Research-funding organisations' index
	Social media outreach/take-up of open access literature
	Research-performing organisations' support structures for researchers as regards incentives and barriers for data sharing
	RRI-related governance mechanisms within research-funding and performing organisations

	RRI-related governance mechanisms within research-funding and performing organisations – composite index

She Figures

There have been several editions of the *She Figures* publication of gender indicators in science, research and innovation in Europe. Several of the indicators presented in She Figures were taken up by the MoRRI project (see Table 2). *She Figures* continues to innovate in the development of new indicators for gender in R&I, for example, in designing an indicator for the gender content of research (Science-Metrix 2015). SUPER MoRRI will consult with the unit responsible for *She Figures* within the EC, with a view to understanding further plans or developments underway with regard to gender indicators and seek ideas on gaps in current monitoring.

Eurobarometer

The Eurobarometer surveys related to public values, perceptions, and attitudes in relation to science and controversial technologies have been carried out since 1978. Numerous surveys have been fielded since, and collectively they offer complex time-series data that shed light on the evolution of the relationship between science and society. Core markers of the social robustness of the interaction of research, innovation, and citizens, such as citizens' trust in scientists, optimism about new technologies, literacy and efficacy in relation to science, and patterns of engagement with science, are captured by the Eurobarometer series.

In the MoRRI project, a couple of indicators, primarily relating to public engagement, were based on Eurobarometer data. It was, however, not updated data, and as some of the main time-series related to aspects of responsibility were interrupted around 2010, the overall relevance of the Eurobarometer as a source of up-to-date information, has been limited. In Spring 2020, however, a new Eurobarometer purposively restoring some of the main time-series will be fielded, and the SUPER MoRRI project will be able to exploit this important resource for building improved indicators, particularly for public engagement, science literacy and education, and open access.

Open Science Monitor

The Open Science Monitor (OSM) was commissioned by the EC to support open science in Europe. It provides data and information, including indicators, on open access publishing, open research data and open forms of collaboration, including citizen science. Openness in science is here conceived as far broader than the definition of open access as a key area of RRI. Identifying and potentially building on relevant, feasible and useful elements for monitoring open knowledge practices is important for the SUPER MoRRI framework. SUPER MoRRI will therefore review the metrics and outputs developed

by the OSM and consult with its developers as part of its development strategy in relation to open science.

3.4.2   Indicators in the Lab

There are also relevant indicators, or suggestions for indicators, that remain within the confines of the laboratory. SUPER MoRRI will consider as many of these contributions as possible and continue to monitor emerging indicators over the period of the project. The potential of two particularly relevant sets of indicators in the lab are summarised briefly below.

Expert Group for Policy Indicators for RRI 2015

The potential indicators outlined in this report were described as forming a toolbox that could support policy development along the lines of individual RRI key areas. Indicators of three types, process, outcome and perception indicators. No attempt to operationalise the indicators proposed have occurred to date. The framing of the Expert Group's policy indicators placed a strong emphasis on taking responsible and constructive approach to indicator design and use, providing insights which are of direct relevance to SUPER MoRRI.

There are three aspects of particular interest in this report that could be taken up in SUPER MoRRI. First, while (as described above) significant progress has been made on designing and populating indicators for the gender and open access/science keys of RRI, other keys conspicuously lack monitoring tools. Second, the Expert Group proposes indicators of different types to those currently in use for gender and open access/science that may add richness to monitoring those keys. Third, the Expert Group proposed the addition of two additional keys, sustainability and social justice/inclusion, which may be of interest for monitoring to support RRI.

In relation to sustainability, for example, the relevant question for monitoring was defined as "to what extent does a research field, a research programme or an RRI initiative contribute to sustainable growth?" (Strand et al. 2015: 7). With sustainability being a pressing question that confronts science and technology in seemingly ever more urgent ways, the Expert Group suggestion, or variant of it, seems ripe for inclusion in a monitoring framework for RRI. Potentially de-linking the monitoring of sustainability in R&I from the question of economic growth might increase the opportunities for relevant and useful monitoring, whilst reducing the difficulties of designing a feasible data and information gathering approach.

Expert Group for Researchers' Engagement with Open Science 2019

A recent contribution to policy and indicator development in the domain of open science is the report of the EC Expert Group for researcher engagement with OS. This report proposes a series of four "indicator toolboxes" to support four open science policy goals:

- Infrastructure indicators oriented to the scientific system at national, international and disciplinary levels;
- Indicators of open knowledge capabilities in research communities;
- Indicators of pioneering open knowledge practices; and
- Individual level indicators for careers (Wouters et al. 2019: 17).

A compilation table of 149 potential indicators for these toolboxes is provided, sourced from organisations, groups and individual researchers concerned to support the development of open science practices and culture. The Expert Group contend that these indicators should only be used in the context of appropriate toolbox frameworks as there are "no generally valid open science indicators that can capture the diversity of open knowledge practices" (Wouters et al. 2019: 16). The indicators are compiled with relevant information on strengths, weaknesses, potential and risks, as well as their current availability. This provides a valuable resource for assessing which of these indicators might be relevant and useful for SUPER MoRRI. It is also a useful model for working with and communicating potential indicators for RRI monitoring to a wider audience.

3.5   Country Correspondents Network

The Country Correspondents Network (CCN) is composed of one correspondent from each of the European member states and selected additional countries. Individual correspondents have been appointed based on their knowledge of the national research and innovation system, their expertise in matters related to RRI, and their ability to carry out data collection rigorously and efficiently. The CCN constitutes an important resource within the monitoring framework, as it provides the opportunity for carrying out comparative, qualitative research with very broad geographical coverage, in a way that is culturally sensitive and based on deep knowledge of the particularities of the respective country contexts. In effect, the CCN will be contributing to the strategic principle of credible contextualisation in relation to the empirical research programme of SUPER MoRRI. For seven countries represented in the consortium (Austria, Denmark, Germany, The Netherlands, Norway, Spain and the UK), partner organisations will function as correspondents, while for the remaining EU countries, country correspondents have been recruited through an open call.

It is the intention that the CCN will facilitate data collection related to both patterns studies and process studies (see below). Specifically, the CCN will enable qualitative studies of research performing and research funding organisations, combining desk research, document analysis, and interviewing. To ensure consistency and comparability, and as an element of quality assurance in relation to preparing data collection exercises, the CCN will gather at a workshop in Vienna on March 26-27, 2020. The

purpose of the workshop is to test study protocols and interview guides, and to arrive at a harmonised understanding of the tasks.

3.6   International Satellite Partners

The International Satellite Partners (ISP) will provide an international dimension to work on monitoring to support RRI. ISPs from Argentina, Australia, Brasil, Canada, China, Iran, Japan, South Africa and the USA will be involved in discussions to develop a set of recommendations concerning the relevance and feasibility of monitoring RRI, particularly in non-European contexts. Some ISPs will also examine the nature of potential data and information sources in their countries. Collaboration with ISPs will be done through a series of interviews and web-based consultations. In 2019, the ISPs provided a set of written critical comments and feedback on the SUPER MoRRI Briefing Paper on the purposes and needs of organisations for RRI monitoring from a global perspective (deliverable 4.1).

3.7   The RRI Project Ecosystem

A rich array of RRI projects and partners exist who can contribute and benefit from collaborating on the development of a monitoring framework for RRI. The project RRI ecosystem aims to connect participants from H2020 SwafS funded projects concerned with RRI project implementation and monitoring issues. The ecosystem has already been established with participants from the FIT4RRI, SHERPA, SeeRRI, Siscode, I AM RRI and On-Merrit projects. The ecosystem will remain open and anticipates growing to include representatives from the many SwafS projects that have either just started or will do so in 2020.

The initial phase of the project ecosystem consists of a virtual meeting space, which convenes monthly. The first meeting was primarily about setting up the ecosystem, understanding the needs and expectations of participants about what the ecosystem could provide, and with introductions and discussions about each participants' RRI projects. Discussions quickly transitioned to some common concerns that were felt within the RRI practice environment. In particular, issues related to language and translation were shared and strategies were identified, within the discussion group, for overcoming translation errors when communicating with diverse populations regarding RRI issues (co-developing surveys and tools through conversations with diverse stakeholders, for example). This was immediate evidence that a conversation space for working on common concerns would be one benefit flowing from the ecosystem concept.

Ambitions for the project ecosystem's future include functioning as a responsive interactive focal point for sharing RRI-related resources that have been created through the diversity of RRI projects. A need was also identified to work collectively on a common dictionary and handbook for RRI resources and concepts that would be usable across linguistic, disciplinary, and sectoral boundaries. In the future it

is anticipated that the project ecosystem could facilitate pilot interviews or focus groups, or act as a test-bed for potential quantifications and tools under consideration within the SUPER MoRRI project. For the ecosystem to succeed in these ambitions, it will require a sustained effort to build productive interactions among RRI project participants, including SUPER MoRRI.

3.8   Self-Assessment Tool

An online assessment method will be developed that can be of assistance to all types of stakeholder organisations to plan and review their own RRI policies and practices over time. It might also be used by them as a background to compare their responsibility profile against similar organisations. Different types of stakeholders will be encouraged to tailor their approach to suit their institutional culture, their operational context and objectives, and the content of their RRI strategy and commitments. Each stakeholder will be able to use the dashboard to better understand what RRI means in practice, by looking at existing examples. The self-assessment tool could therefore also serve to supplement policy and practice learning about RRI, whilst potentially building a new user-driven data ecosystem centred on innovation.

4. RESEARCH PROGRAMME

SUPER MoRRI is, above all else, a research project. We will be designing and conducting a number of research studies that will generate primary data and link to and use secondary data from various existing sources. The SUPER MoRRI research programme will be managed through two thematic streams: Pattern Studies and Pathway Studies. While we will have a set of primary data collection vehicles underpinning our pattern studies, they will not be isolated but will contribute to addressing research questions that draw on several of these vehicles. Pathways studies will explore issues identified in exploiting these data vehicles. Whilst some studies will be designed either to continue or establish potential time-series data collections, others will drive toward identification of issues and opportunities or toward deepening our understanding. A first wave of studies will be initiated from early 2020, followed by a second wave from early 2022 (Figure 1). In addition, a process of reflection and assessment will take place in late 2021, at which time quality assessments will be made, emerging issues clarified and, if desirable and feasible, a smaller second wave of studies will be designed to take advantage of new opportunities.

4.1 Pattern Studies

A number of different data vehicles will be developed that provide the opportunity to analyse patterns of different sorts, including geo-spatial, disciplinary or organisational comparisons. A range of quantitative and qualitative data sources will be used to develop new data and information that can complement these data vehicles. Empirical studies are in the planning phase that will utilise these data vehicles related to research funding organisations, research performing organisations, citizens and researchers/research groups. Details of the role of data vehicles in our empirical work, including their key role in pattern studies, are set out in the accompanying SUPER MoRRI Implementation Plan (deliverable D2.1).

4.2 Pathway Studies

A number of research projects will focus on improving our understanding of processes and practices related to the institutionalisation of responsibility in research and innovation. The studies will aim to highlight how diverse sets of actors are contributing, individually and collectively to these efforts, including by:
- institutionalising responsible knowledge practices (Randles 2017), taking into account the inequality in distribution of resources, infrastructures and capabilities in research and innovation communities (Wouters et al. 2019);
- building capabilities for responsible practices in scholarly knowledge production and peer communities; and

● transforming incentives, rewards, evaluation, promotion and other criteria to reduce barriers to institutionalising responsible practices and values (Owen and Pansera 2019).

These studies will provide opportunities for appropriate comparisons and the identification of exemplar cases. An important contribution of pathways studies for monitoring to support RRI will be to provide relevant peer-comparisons at the level of processes and practices and to point toward exemplary models and initiatives that can speed up learning.

The process for designing and selecting SUPER MoRRI pathways studies and details on their research approach and timing are set out in a separate Case Research Plan (deliverable D5.1).

4.3   Secondary Data Sources

A number of secondary data sources will play important roles in the empirical programme. Some will be vehicles for quantifications, including indicators. Others will be sources of secondary data that will support interpretive models and complementary information that will support users of the monitoring framework. Important sources of secondary data include EUROSTAT, Eurobarometer, the European Tertiary Education Register (ETER), PATSTAT and bibliometric databases, among others.

Figure 2: Data vehicles for SUPER MoRRI
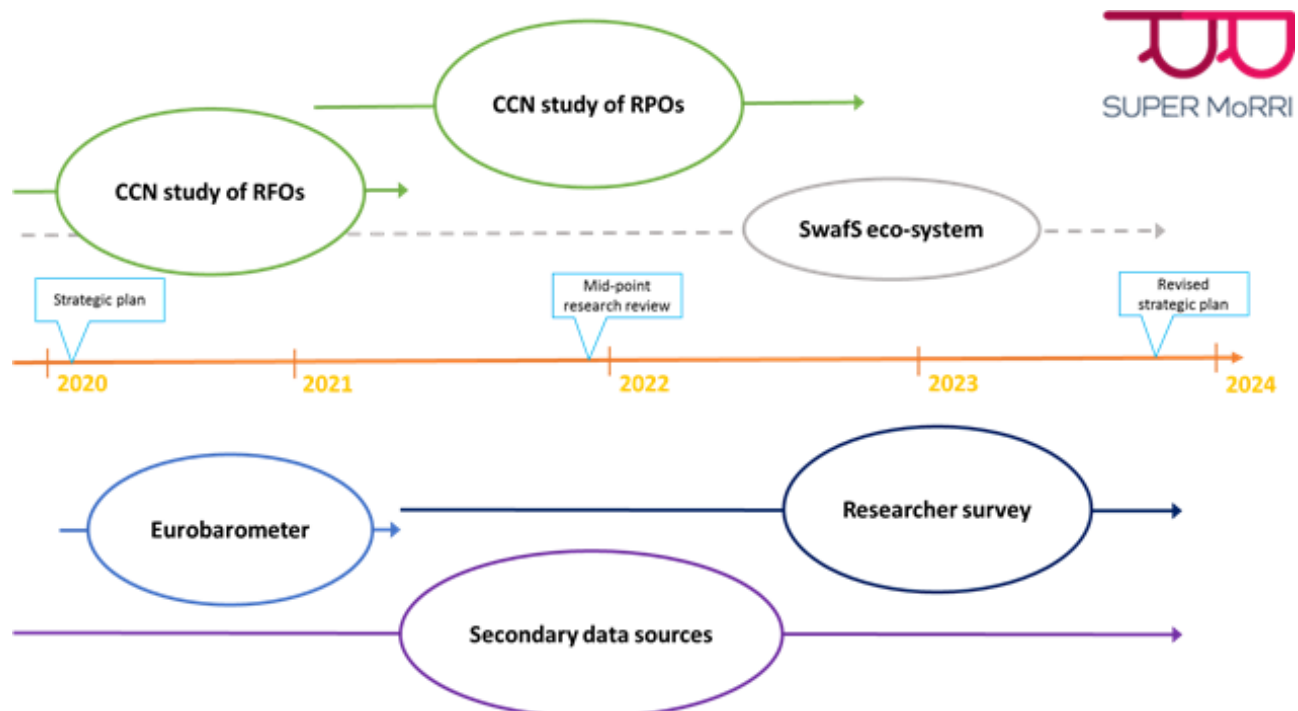
4.4   Monitoring Reports

SUPER MoRRI plans to deliver three Monitoring Reports during the life of the project in December 2020, April 2022 and August 2023. Each Monitoring Report will add new elements as they become available from the empirical research programme or from identification of other relevant secondary sources. Details can be found in the Implementation Plan.

4.5   Indicator Co-creation

In the case of the development of new indicators, a co-creation phase will be conducted. This will involve bringing together a small number of potential users of the indicator to critically reflect on the work-in-progress and offer their advice. This process will be iterative and involve both co-creation of the final indicator and the form in which it will be presented. Presentation of new indicators will include three components: an **indicator fiche**; a suggested **interpretive model**; and supporting **complementary information**.

The indicator fiche will include all relevant technical information, including the data source(s), the metric used to calculate the score, and the indicator coverage. Data used will be treated in accordance with the SUPER MoRRI Data Management Plan (deliverable D8.1).

Suggested interpretive models for indicators will be tailored to each indicator. In the case of one-off indicators the interpretive model will explain the rationale for the creation of the indicator and how it is perceived to support RRI. For indicators that are time-series, or have the potential for future replication to create time-series, the model will describe what a change in the indicator can be reasonably understood to mean.

Complementary information will be provided to try and ensure the credible contextualisation of the indicator. Complementary information will include descriptive information on specific conditions, such as regional or national contexts or stages of policy cycles that will support users in understanding the limits of what the indicator can be thought to validly and reliably signify.

The combination of these three elements is intended to ensure that resources are available to support users to make appropriate indicator-based comparisons.

5. EMBRACING STRATEGIC UNCERTAINTY

The transition from Horizon 2020 to Horizon Europe, the next European Framework Programme for research and innovation investment programme (2021-2027), means changes to the policy context for RRI. From the perspective of planning and developing a monitoring framework to support responsibility in research and innovation, the effects of this policy transition cannot be entirely foreseen. Nevertheless, the discontinuation of the SwafS work programme in Horizon Europe is known and is being taken into account in SUPER MoRRI. One effect of this change is that no further SwafS projects will come online after Horizon 2020 is finalised. This will impact on the capacity to build a RRI-project-based data ecosystem into the medium-term future.

Whilst the position of RRI, as a policy discourse, set of instruments, projects and a growing community of practitioners, is in transition, SUPER MoRRI takes the view that responsibility in research and innovation overflows any specific framing in theory, policy or otherwise. The strategic vision of SUPER MoRRI is o provide a monitoring framework and set of user tools that will have have value into the future. This vision includes both those stakeholders who may consider themselves to be tied quite closely to RRI and its implementation, and those with a broad interest in understanding and influencing the direction of travel of the complex of relationships among science, research, innovation and society.

REFERENCES

Callon, M., P. Lascoumes & Y. Barthe (2009) *Acting in and Uncertain World: An Essay on Technical Democracy*. Cambridge MA, MIT Press.

European Commission (2014) Rome Declaration on Responsible Research and Innovation in Europe. Rome, European Commission.

European Commission (2017) EN Horizon 2020 Work Programme 2018-2020 16. Science with and for Society. Brussels, European Commission.

European Commission (2019) *She Figures 2018*. Directorate-General for Research and Innovation.

Fisher, E. (2018). Ends of responsible innovation. *Journal of Responsible Innovation 5*(3): 253-256. https://doi.org/10.1080/23299460.2018.1513900

Funtowicz, S. & J. Ravetz (1993) Science for the post-normal age. *Futures* 25(7): 739-755.

Hicks, D., P. Wouters, L. Waltmann, S. de RIjke & I. Ràfols (2015) Bibliometrics: The Leiden Manifesto for research metrics, *Nature* 520, 429–431.

Hormia-Poutanen, K., E. Kristiansen, R. Lawrence, S. Leonelli, N. Manola, E. Méndez, C. Rossel, M. Vignoli & M. Dias Agostinho (2017) 'Recommendations of the OSPP on Next-Generation Metrics', Altmetrics Working Group of the Open Science Policy Platform, Recommendations of the OSPP on Next-Generation Metrics, adopted on 30 October 2017

Kleibrink, A., C. Gianelle & M. Doussineau (2016) Monitoring innovation and territorial development in Europe: emergent strategic management, European Planning Studies 24:8, 1438-1458, https://doi.org/10.1080/09654313.2016.1181717

Lehtonen, M. (2017) Operationalizing information: measures and indicators in policy formation, Ch. 10 in M. Howlett & I. Mukherjee (eds.) Handbook of Policy Formation, Elgar Online, pp. 161-181, https://doi.org/10.4337/9781784719326.00017

Lindner, R., S. Daimer, B. Beckert, N. Heyen, J. Koehler, B. Teufel, P. Warnke & S. Wydra (2016) Addressing directionality: Orientation failure and the systems of innovation heuristic. Towards reflexive governance. Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis No. 52, ISSN 1612-1430 http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-4087463.pdf

Mejlgaard, N., R. Woolley, C. Bloch, S. Bührer, E. Griessler, A. Jäger, R. Lindner, E. Madsen, F. Maier, I. Meijer, V. Peter, J. Stilgoe & M. Wuketich (2018) A key moment for European science policy. *Journal of Science Communication*, 17(3) https://doi.org/10.22323/2.17030305

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

Owen, R. & M. Pansera (2019) Responsible Innovation and Responsible Research and Innovation. Chapter 2 in D. Simon, S. Kuhlmann, J. Stamm & W. Canzler (eds.) Handbook on Science and Public Policy. Cheltenham, Edward Elgar: 26-48.

Peter, V., F. Maier, N. Mejlgaard, C. Bloch, E. Madsen, E. Griessler, M. Wutekich, I. Meijer, R. Woolley, R. Lindner, S. Bührer, A. Jäger, L. Tsipouri & J. Stilgoe (2018) Monitoring the evolution and benefits of responsible research and innovation in Europe, Final Report of the MoRRI Project. Directorate-General for Research and Innovation, Open Innovation and Open Science, Brussels. Final report – Summarising insights from the MoRRI project

Ràfols, I. (2018) S&T indicators in the wild: Contextualization and participation for responsible metrics. *Research Evaluation* 28(1): 7-22.

Randles, S. (2017) Deepening 'Deep Institutionalisation'. JERRI project deliverable 1.2. https://www.jerri-project.eu/jerri-wAssets/docs/deliverables/wp-1/JERRI_Deliverable_D1_2_Deepening-Deep-Institutionalisation.pdf

Rommetveit, K., N. Dijk, K. Gunnarsdottír, K. O'Riordan, S. Gutwirth, R. Strand & B. Wynne (2019) Working responsibly across boundaries? Some practical and theoretical lessons. Chapter 6 in R. Von Schomberg (ed.) *International Handbook of Responsible Innovation*. Cheltenham, Edward Elgar: 83-100.

Science-Metrix (2015) She Figures 2015; Comprehensive Methodology – New Research & Innovation Output Indicators. Brussels, Directorate-General for Research and Innovation. http://www.science-metrix.com/sites/default/files/science-metrix/publications/she_figures_2015_comprehensive_methods_on_ri_outputs_final.pdf

Stilgoe, J., R. Owen & P. Macnaghten (2013) Developing a framework for responsible innovation. *Research Policy* 42(9): 1568–1580.

Stirling, A. (2008) "Opening Up" and "Closing Down": Power, Participation, and Pluralism in the Social Appraisal of Technology. *Science, Technology, & Human Values*, 33(2): 262–294. https://doi.org/10.1177/0162243907311265

Strand, R. J. Spaapen, M. Bauer, E. Hogan, G. Revuelta, S. Stagl, L. Paula, Â. Guimarães Pereira (2015) Indicators for promoting and monitoring Responsible Research and Innovation. Report from the Expert Group on Policy Indicators for Responsible Research and Innovation. Brussels, Directorate-General for Research and Innovation.

Strand, R. & K. Rommetveit (2019) SUPER MoRRI Concept Note, University of Bergen (UiB), March, pp. 1-10.

Timmermans, J. & V. Blok (2018) A critical hermeneutic reflection on the paradigm-level assumptions underlying responsible innovation. *Synthese* https://doi.org/10.1007/s11229-018-1839-z

Von Schomberg, R. (2013). A vision of responsible research and innovation. In R. Owen, M. Heintz, & J. Bessant (Eds.) *Responsible innovation: Managing the responsible emergence of science and innovation in society*. Chichester, Wiley: 51–74,

Von Schomberg, R. (2014) The quest for the 'right' impacts of science and technology: A framework for responsible research and innovation. In M. van den Hoven, N. Doorn, T. Swierstra, B. Koops & H. Romijn (Eds.) *Responsible Innovation 1: Innovative Solutions for Global Issues*. New York, Springer: 33-50.

Von Schomberg, R. (2019) Why Responsible Innovation? in R. Von Schomberg & J.Hankins (Eds.) *The International Handbook on Responsible Innovation. A Global Resource*. Cheltenham and Northampton, Edward Elgar: 12-32.

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

Wilsdon, J., J. Bar-Ilan, R. Frodeman, E. Lex, I. Peters, & P. F. Wouters (2017). Next-Generation Metrics: Responsible Metrics and Evaluation for Open Science. Report of the European Commission Expert Group on Altmetrics. European Commission, Brussels. https://ec.europa.eu/research/openscience/pdf/report.pdf

Wouters, P., I. Ràfols, A. Oancea, S. C. L. Kamerlin, J. B. Holbrook & M. Jacob (2019) Indicator frameworks for fostering open knowledge practices in science and scholarship. Report of the Expert Group on Indicators for Researchers' Engagement with Open Science. Brussels, Directorate-General for Research and Innovation.

Appendix

Task 1.2 Working Paper: Critical Inspection of MoRRI Indicators

Working Paper: Task 1.2 – Critical Assessment of the Existing MoRRI Indicators and Monitoring System

Author(s): Thomas Kjeldager Ryan (AU) & Niels Mejlgaard (AU)

Table of Contents

1 Introduction

1.1 Scope and objectives of the deliverable

The aim of task 1.2 is to perform a critical assessment of the existing MoRRI indicators and monitoring system. We assess the validity, feasibility and relevance of the MoRRI indicators. The Monitoring system is assessed in terms of how well the set of indicators reflect the RRI keys and sub-concepts and the overall feasibility of gathering primary and secondary data in the context of SUPER MoRRI.

The MoRRI project compiled 42 national level indicators in order to report on the state of affairs of RRI in EU. Each indicator relates to one of the six keys of RRI: Gender Equality (GE), Science Education and Science Literacy (SLSE), Public Engagement (PE), Open Access (OA), Ethics and Governance (GOV). The MoRRi project assessed the indicators in terms of robustness, relevance, and richness. Prior validity testing assessed the empirical validity of the indicators, mainly in terms of sensitivity analysis. In the cases of composite indicators, Cronbach's alpha was used to assess the empirical validity of composites. Finally, comparison of variation between and within countries was performed to assess the validity of aggregating meso-level indicators to the national level (Intraclass variation).

The current assessment is based on the documents produced in MoRRI. Namely, the technical appendix (Appendix_D4.3_20022018_clean) as well as the data files complied in MoRRI (Masterfile D4.3).

The components of the critical assessment are:

- Concept-indicator congruency or (Content validity): What is the aim of the indicator in terms of the concept it seeks to represent? To what extent is the developed indicator and the overall concept congruent?
- Relevance (usefulness, actionable, reflective): Is the indicator useful and actionable for stakeholders i.e. does it allow for reflection on the overall concept?
- Technical validity – (Reliability): Is the data for the indicator collected in the most appropriate way, and in the most effective manner? Can the question be answered consistently depending on the respondent or data collection procedure?
- External validity: Is the quality of the data collected in MoRRI sufficient to generalize and at what level? What would a minimum threshold of responses/observations be to discuss between case and time comparisons?
- Future data collection feasibility: To what extent is future data collection to populate the indicator plausible?

1.2 Structure of the deliverable

1) Overview and background of MoRRI
2) Critical evaluation of indicators based on primary data
3) Critical evaluation of indicators based on secondary data
4) Summary of indicator evaluation using a colour coding system: A green, yellow or red code is given overall and in terms of relevance, validity and feasibility.
   a) Green: The indicator can be collected without any major improvements or resources

b) Yellow: There are issues with relevance, validity or feasibility that need to be addressed before collecting data.

c) Red: There are significant constraints that make data collection and indicator construction very difficult or unfeasible.

5) Comments on challenges and suggestions for SUPER MoRRI.

1.3   Relation to Other Tasks and Deliverables

The assessment provided in this document will inform task 1.8, the definition of a continuing MoRRI indicator set. The indicators that are deemed relevant, feasible and of a high validity will be candidates to enter into this final set. Furthermore, possible new indicators proposed in this document will be considered in task 1.8. Additionally, the assessment of the indicators and monitoring system will be used in task 1.9 as a source of knowledge for producing revised data fiches and to spell out the purpose of the monitoring system as a whole as well as the individual indicators.

1.4   Purpose & Scope of indicators in MoRRI

The MoRRI indicators had the purpose of providing actionable and comparable information on the status of RRI in EU. In MoRRI, 42 national level indicators were constructed and compiled. The process of compiling the final set of indicators was based on five reports on the six RRI keys that provided a conceptualization of each key and a mapping of potential indicators. The mapping resulted in a list of 97 potential indicators. MoRRI found 42 to be feasible, relevant and valid.

The number of indicators under each key varies, so does the data source. Many of the indicators are based on primary data collection through surveys, while the remaining are secondary data, which often already are used as indicators in related settings (E.g. Eurobarometer and She-figures).

RRI Key	Indicators (Primary data)
Gender Equality	15 (5)
Science Education and Science Literacy	4 (2)
Public Engagement	10 (5)
Ethics	5 (4)
Open Access	6 (4)
Governance	3 (2)

1.5   Data Sources for MoRRI indicators & indicator overview

In this section, we shortly provide an overview of the primary and secondary data sources, as well as the features of the empirical material collected (e.g. response rates and absolute responses, level of data-collection etc.). Primary data sources include: The public research organisation (PRO) survey and the higher Education (HEI) Survey (When combined the two surveys are referred to as the Research performing organisation (RPO) survey). The research funding organisation (RFO) survey and science in society actors (SiS) survey as well as data generated by country correspondents. Secondary sources

include: Bibliometric databases (Scopus, Altmetric.com & Patstat), Eurobarometer, She-figures, Eurostat and MASIS Relation to Other Tasks and Deliverables.

1.6   RRI Concepts

The conceptualization of each key in the MoRRI project is briefly summarized below, using excerpts from the MoRRI outputs.

1.6.1   Gender Equality

"[…] in the context of the MoRRI project, we understand gender equality as a three-dimensional construct. To achieve gender equality, progress is required on all three dimensions. Accordingly, the three equality goals are:

1) Integration of women in all fields and at all levels in research and innovation (reduction of horizontal and vertical segregation

2) Structural change in research institutions in order to abolish structural barriers for women (e.g. through implementation of comprehensive equality plans, quotas for women, transparent decision-making), and

3) Integration of gender in research and innovation content to ensure that the needs and interests of women are adequately addressed." [2171-D2_3_Gender_Equality, page 19]

1.6.2   Science Literacy and Science Education

"[In] this project [MoRRI] we define science literacy as the ability of citizens to read about, comprehend and express opinions 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 can be generated through three main mechanisms:

1) 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.

2) 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.

3) Co-production of knowledge is characterised by a co-creation of knowledge through cooperation of scientific experts and non-experts." [2172-D2_2-Science_Literacy p.2]

1.6.3   Public Engagement

"The five categories below are identified:

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 in order 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 characterized 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)."
[2171-D2_1-Public Engagement]

1.6.4 Open Access

"Open Access (OA): Open access is the idea of making research results freely available to anyone that wants to access and re-use them. One of the main drivers of the OA idea is to make publicly funded research accessible to the general public. In the academic sense, the term Open Access referred originally to the provision of free access to peer-reviewed academic publications.

Open Data (OD): Presently, the term [open access] also encompasses the free access to the research data that underpins publications or research projects, also referred on its own as Open Data (OD). Open Data is usually distributed with requirements of attribution and share-alike (copies or adaptations of the data need to be shared using the same principles as the source)." [D3.1.-final]

1.6.5 Ethics

"The authors draw useful distinctions on how to delineate the institutionalisation of ethics and how to categorise in

ethical governance, i.e. "institutionalizing ethics debate in terms of the implementation of standards in research ethics in science, technology and innovation policies" (Brom et al., 2015, p. 15);

ethical deliberation, i.e. "institutionalizing ethics debate that raise issues in science and technological developments in science, technology and innovation policies" (ibid.);

ethical reflection, i.e. "institutionalizing 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" (ibid.)." [2171-D2_4.1-Ethics]

1.6.6 Governance

"Governance is defined as steering innovation through the establishment of goals, the establishment of means and the verification of performance. For science and innovation, this therefore means the provision and distribution of funding and the regulation of research and innovation activity through soft (cultural and normative) means as well as hard ones (laws and institutional procedures). The governance of science is concerned with how knowledge is produced and how it is distributed. The governance of innovation is far broader. 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." [2172-D2_4.2 Governance]

2. INDICATOR ASSESSMENT

The Indicator assessment is presented in two parts. The first assesses indicators based on primary data and the second on secondary data. This has been done, because the assessment is different depending on whether we have influence on how data is collected and treated.

2.1 Assessment of primary data

2.1.1 GE1 - Share of RPOs with gender equality plans

Measurement and definition

GE1 is based on question 19 in the PRO and HEI survey.

> "Does your organisation have a gender equality plan? A gender equality plan is a consistent set of provisions and actions aimed at ensuring gender equality".

The possible responses are, "Yes", "No", "Not Applicable" and "Do not know".

The indicator is calculated as the share of institutions in a country answering "yes" out of all valid responses ("Yes", "No" or "not applicable").

Congruency and relevance

The indicator indicates the extent to which gender equality is institutionalized through local organisational policy initiatives. In this sense the indicator represents the concept *structural change in institutions* which is described in the section above.

The indicator is relevant for policy-makers, as it allows for an overall view of the degree to which gender equality is institutionalized. For higher education institutions and research performing organizations the indicator provides an opportunity to evaluate whether they are a part of the majority or minority in their country and Europe in general.

Validity

The statistical test performed in MoRRi was Interclass correlation (0.47). The test indicates a high share of variation is between countries relative to within country [Appendix D3.4]

The responses to the question indicate that most of Higher education institutions have a gender equality plan, and less than one-third PROs indicate that they have a gender equality plan (See table GE1). Therefore, the country level indicator is sensitive to the sampling of HEIs and PROs in a country.

Table GE1

type	No	Not app..	Not known	Yes	Total
HEI	40	65	16	22	111	254
	15.75	25.59	6.30	8.66	43.70	100.00
PRO	33	77	20	11	67	208
	15.87	37.02	9.62	5.29	32.21	100.00
Total	73	142	36	33	178	462
	15.80	30.74	7.79	7.14	38.53	100.00

A challenge for this indicator is that it cannot provide a grading of the degree of gender equality focus in organizations. Another type of questions is needed to capture degrees of GE focus and actions in RPOs. The indicator is a country-level indicator based on organisational level data. The external validity is therefore contingent on the response rates and absolute responses in each country.

Evaluation of validity, relevance and feasibility

The indicator represents the degree to which gender equality is institutionalized in the member states (structural change in institutions). It is highly relevant to policy-makers to assess the MSs Practices and actions within Europe. However, it only captures a small part of the overall concept and cannot alone be used to discuss the institutionalization of gender equality in Europe (structural change in institutions). PROs and HEIs differ in terms of the proportion that have a GE plan. Therefore, the sampling of HEIs and PROs within a country is a relevant parameter to consider.

The summary evaluation of the indicator is green. Changes to the collection of data and calculation of the indicator are not necessary. However, the level at which the indicator is reported can be discussed. The weaknesses of the indicator relate to the aggregation to the Member state level.

She-Figures also reports on the same question, therefore, it can be possible to use this secondary source as an alternative.

2.1.2   GE3: Share of RFOs promoting gender content in research

Measurement and definition,

GE3 is based on question 20 in the RFO survey.

> "When allocating research and innovation funding in years 2014, 2015 and 2016, did your organisation include the gender dimension in research content?"

The possible responses are: "Yes, standard criterion in all programmes", "Yes, standard criterion in specific types of programmes", "No/ Not App" , "Do not know"

In the construction of the indicator, "Yes, standard criterion in all programmes" is given a value of one and "Yes, standard criterion in specific types of programmes" is given a value of a half and no or not applicable is given a weight of zero. The national level indicator is the average score for responding organisations in each country.

Congruency and relevance

GE3 reflects the degree to which assessments of proposals and thus access to research funding is dependent on researchers taking gender and sex analysis into consideration. Thus, there is congruency between the sub-concept "Gender in research content" and the indicator.

The indicator is relevant for both university administrators and policy makers to be able to follow the degree gender and sex analysis is a specific or standard criterion in the national funding systems.

Validity

In the MoRRI report two tests were performed: Intraclass variation (0.42) and the effect on ranking by changing the indicator to a binary variable. The conclusion is that there was more variation between than within countries and that the indicator did not change the ranking of countries considerably.

The results show that 31 out of 106 organisations use "the gender dimension" as a criterion in evaluation to some extent (table GE3). The, question thus succeeds in distinguishing between practices in research funding organisations. It is not explicitly stated that the question regards gender and sex analysis, therefor the validity of the indicator is reduced.

The calculation of the indicator assumes a value of 0.5 for having a standard criterion in specific programs versus a value of 1 for a general criterion in all programs. This can be argued to be somewhat arbitrary. In the MoRRI report the effect of changing the indicator to a binary variable was found to be low. Therefore, an option is to simplify the calculation of the indicator and thus avoid the choice of assigning values to the two confirmatory responses.

Table GE3

19.3. 2016	Freq.	Percent	Cum.
	18	16.98	16.98
Don't know	6	5.66	22.64
No	51	48.11	70.75
Yes, it was a standard criterion in a..	18	16.98	87.74
Yes, it was a standard criterion in s..	13	12.26	100.00
Total	106	100.00	

Evaluation of validity, relevance and feasibility

The indicator measures the third pillar within gender equality. The indicator is limited by the lack of a definition of "the gender dimension", which creates validity issues. The calculation of the indicator could be simplified to a binary code.

The summary evaluation of the indicator is yellow: a good indicator yet, a definition of "the gender dimension" would improve the validity. Second, aggregation of the indicator to the Member state level is associated with external validity issues because of the low number of absolute responses in each country.

2.1.3 GE5: Share of RPOs with policies to promote gender in research content

Measurement and definition

The indicator is based on responses to Q20 in the PRO and HEI survey.

> Does your organisation have implemented processes to promote the integration of a gender dimension in research and innovation content of projects and studies, for example information and qualification tools or concrete rewards and incentives?

The possible responses were: "Yes", "No", "Not Applicable" and "Do not know".
The indicator is calculated as the share of institutions in a country answering "yes" to implementing processes out of the total number of responses ("Yes", "No" or "not applicable").

Congruency and relevance

GE5 relates to with the overall concept of "gender in research content" and in part with "structural change in institutions". It reflects the degree to which institutions in the MSs have attempted to institutionalize research practices that include gender and sex analysis in research.

GE5 is a relevant indicator that provides information on the third pillar of gender equality from the perspective of organisations. It compliments GE3 that assesses the third pillar from the perspective of funders. The combination of the two indicators can provide an understanding of how prevalent the third pillar is in the system in general.

Validity

The question does not define the concept "a gender dimension". The lack of this definition creates issues of interpretation.

In a following question, the respondents had the option to indicate the types of initiatives they have in place du exemplify why they answered yes. The various descriptions show how respondents understood the question very differently. Some understood the question to pertain to gender equality in an employment equality sense, while few found it to relate to an emphasis on gender in research content and mentioned establishment of gender research centres, and a focus on gender and sex analysis.

Second, the grammar is incorrect, and should be re-phrased.

Table GE5.1

type		No	Not app..	Not known	Yes	Total
HEI	39	91	21	31	72	254
	15.35	35.83	8.27	12.20	28.35	100.00
PRO	30	87	29	14	48	208
	14.42	41.83	13.94	6.73	23.08	100.00
Total	69	178	50	45	120	462
	14.94	38.53	10.82	9.74	25.97	100.00

Below three examples from the following question are given to show how respondents understood the concept of "the gender dimension".

21. Previously you indicated that your organisation has implemented processes to promote the integration of a gender dimension in research and innovation content of projects and studies. Could you please briefly describe the processes in place?

1) "Creation of a situation report comparing men and women (notably concerning their research careers (…)"
2) "A dedicated nursery has been established to facilitate the work of women in academia"
3) "All staff are now required to complete an on-line unconscious bias training (…)"

Evaluation of validity, relevance and feasibility

The indicator represents the degree to which gender and sex analysis is institutionalised RPOs. It is highly relevant to policy-makers to assess the MSs Practices and actions within Europe. The indicator is congruent with the underlying concept. However, it only captures a small part of the overall concept and cannot alone be used to discuss the institutionalization of gender equality in Europe.

The summary evaluation of the indicator is yellow: The indicator has some important qualities, and with an addition of a definition of the central concept, it could improve the interpretative validity of the indicator.

An alternative solution could be to pose the question, in a researcher survey, of whether researchers face any incentives that encourage gender and sex analysis in their research. This compliment to or entirely substitute GE3 and GE5.

2.1.4   GE8: Share of female heads of RPOs

Measurement and definition

GE8 is based on Q22 in the PRO and HEI survey:

> Please specify the gender of the person who was/is head of your organisation in 2014, 2015 and 2016 (Head of organisation: highest decision-making official in the organisation (e.g. rector or equivalent in the academy, president or equivalent in non-academic research organisations))

The indicator is calculated as the share of institutions in a country answering "female" relative to total responses.

Congruency and relevance

GE8 provides a partial indication of the vertical participation of women in research. In particular, it provides an insight into the degree to which women are in high ranking managerial positions in higher education and research performing institutions in a given country. This however, should not be interpreted in terms of participation of female researchers in science, as these positions need not have any connection to the gender equality within the research profession. GE8 is a very relevant indicator

for policy makers in order to understand the progress of female participation in management and leadership roles over time and across counties.

Validity

The question is easy to answer, as the information will be readily available even to a person outside the organization.

The phrasing of the question can be improved, in order to minimize the strain on respondents. For example: Please specify the gender of the head of your organization. /Please specify whether the head of your organization is female or male. (Head of organization is the highest…)

The empirical results show that three times as many men as women are heads of RPOs in Europe (see table GE8). There does not seem to be a significant difference between HEIs and PROs. This result reflects the general picture for female participation in high-level jobs in Europe.

Table GE8

type		Female	Male	Total
HEI	40	48	166	254
	15.75	18.90	65.35	100.00
PRO	32	42	134	208
	15.38	20.19	64.42	100.00
Total	72	90	300	462
	15.58	19.48	64.94	100.00

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is green: A simple indicator with a relatively high validity, which could be improved through better response rates and manual data collection. The indicator is closely aligned with the concept of vertical participation in research – although specifically focused on the leadership role. It would be feasible and easily comparable to collect data for future indicator construction.

However, since other indicators of vertical and horizontal participation exist (She-Figures), this indicator provides limited extra value. A possible extension is to show the proportion of female heads of RPOs relative to the proportion of women in high-level jobs in the same country. Such an indicator can show how female participation in higher education is doing relative to other fields in the same country.

Complementary questions regarding lower levels of management i.e. heads of departments or faculties may provide more detailed information. Complementary question regarding the gender composition of

Boards to the organisations may also be an important indicator of female participation and representation in the higher levels of organizations. This information should be relatively easy to collect for a respondent in an organization as it will often be publicly available in an organizational website.

2.1.5   GE9: Share of gender-balanced recruitment committees at RPOs

Measurement and definition

GE9 is calculated on the basis of Q23 & Q24 in the PRO and HEI survey.

> Q23: How many recruitment committees for leading researcher positions did your organisation set up in 2014, 2015 and 2016 for the recruitment of researchers?
> Q24: In how many recruitment committees for leading researcher positions in the share of female members was equal or higher than 40% of the total committee members?

The indicator is calculated as the share of recruitment committees that had at least 40 % representation of women (GE9=Q24/Q23).

Congruency and relevance

This is extremely relevant indicator that provides information on participation of women in important organisational decision-making. Looking at this indicator alongside indicators of vertical workforce participation, can indicate whether gender equality in terms of participation or representation also relates to gender equality in terms of power in decision-making.

The indicator relates both to the vertical participation in science because the availability of more female researchers at a level makes this indicator higher, and in terms of institutionalization of gender equality, because organizations with a high level of female participation regardless of available females in the certain levels may push towards gender balance.

Validity

The indicator is challenged by the difficulty of finding valid information in order to answer the question. Respondents may guess or estimate, which may be influenced by the view and experiences that respondent has (availability heuristic bias). There may also be a risk of social desirability-bias. Finally and most importantly, respondents may choose not to answer because it is too resource demanding to find the information.

Empirically there is a large degree of non-responses as well as a large number of organisations indicating that they had zero committees in the period. Below is the number of Committees in 2016. 201 organisations did not answer the question and 79 answered that they had non in 2016. The large number of non-responses could indicate that the question in fact is very difficult to find data on, and therefore the respondents did not do so.

For those that did respond to both questions and had at least 1 committee there is great variation in the fraction of committees with at least 40 % female. Out of 95 organisations, 38 said that all of their committees had at least 40 % females, while 15 said that non of the committees did.

Evaluation of validity, relevance and feasibility

The indicator is simple to calculate but difficult to respond to because the information will not be readily available for all respondents. Given valid responses and acceptable response rates, the indicator provides relevant information on the degree of "actualized" vertical participation in organisations. As it assesses actions rather than simply head counts of women in the organization it gives a more detailed account of how well females are represented in the decision making in organisations. The indicator complements She-Figures indicators and can be seen in relation to head counts in upper levels of research performing organisations to answer the question of whether women are over or underrepresented in important decision-making activities relative to available women compared to men.

The summary evaluation of the indicator is red: The difficulty of answering and the uncertainty about how precise the responses are. In combination with the low response rate to the question, these issues raise validity questions about the indicator. Because there are a number of good indicators in SHE-Figures that represent similar concepts, the value of collecting data for the indicator again can be discussed.

2.1.6   SLSE1 – Importance of societal aspects of science and science curricula for 15-18 year-old students

Measurement and definition

Country correspondents provided a qualitative assessment based on the following questions.
1) Does the curriculum address the controversial character of either one of the two topics? "yes" "no"
2) Which of the following issues is addressed by the curriculum in relation to the controversial topic (GMO, nuclear energy)?
a. social aspects, such as consequences for the society or agriculture
b. environmental aspects, such as the effects of monocultures or resistances, atomic waste storage etc.
c. ethical aspects, such as development issues like the „golden rice", intergenerational fairness etc.
3) To what degree are they covered? Are they important aspects of the topic or only mentioned in passing? Please briefly explain the reasons for your assessment.

The indicator is calculated by assigning values to the responses and calculating a simple average for each country.

Congruency and relevance

SLSE1 looks at controversial science topics and their coverage in the curricula of 15 to 18-year-old students. This indicator specifically looks at two controversial science topics, genetically modified organisms (GMO) and nuclear energy. It records whether social, economic, environmental and ethical aspects are taught and discussed in relation to these two controversial topics

The indicator relates to the concept of Science education for the youth in a country. There is a high congruency between the concept and the measure. The question is very focused on controversial topics,

therefore, the responses relate to a specific part of science education, where general science education is not directly measured.

Validity

The validity of the indicator rests on the consistency of coding between the country correspondents. Given a thorough methodology the indicator has no obvious validity issues. In future rounds of data collection specific attention should be given to the education and instruction of country correspondents. Here an emphasis on maintaining a similar approach to the first round is imperative in order to ensure comparability over time.

Evaluation of validity, relevance and feasibility

The overall evaluation given is "Yellow": The indicator provides a view into the degree to which the youth in a country is provided with science education. If the indicator is reproduced it would demand a set-up similar to the first round, with country correspondents. For time comparisons, it would demand that the original material in the first round is purposefully used in the instruction of country correspondents. Still, the issue of consistency over time and between countries will be an important focus area.

2.1.7 SLSE2: RRI-related training at HEIs

Measurement and definition

SLSE2 is based on Q25 in the HEI survey:

> Did PhD students' trainings include RRI-related aspects (such as ethical, economic, environmental, legal and social aspects) in 2014, 2015 and 2016?

The possible responses were, "Yes, training in these aspects is mandatory"; "Yes, but training in these aspects is voluntary"; "No", "Do not know" and "Not applicable".
The indicator is calculated by assigning weights to the mutually exclusive categories. "Yes, Mandatory" is given a weight of one and "Yes, Voluntary" a weight of a half and no or not applicable a weight of zero.

Validity

The MoRRi report tested the effect of changing the indicator by weighting both yes answers as one. The effect was that five countries change 5 or more spots in ranking for this alternative.

The main technical issue is the weighting of the categories. The choice of 1 and 0.5 is arbitrary. In addition, the difference between mandatory and voluntary may be difficult to answer. Simplifying the question would provide less interpretive noise and more consistency.

Since a large share respond "not applicable" there may be a number of the HEIs that do not educate PhDs. Second, only 21 respond "no" (see table SLSE2). The question is only given to HEIs. RPOs did not receive the question, probably because they are assumed to not have PhDs very often.

Table SLSE2

25.3. 2016	Freq.	Percent	Cum.
	47	18.50	18.50
Don't know	25	9.84	28.35
No	21	8.27	36.61
Not applicable	35	13.78	50.39
Yes, but training in these aspects is..	72	28.35	78.74
Yes, training in these aspects is man..	54	21.26	100.00
Total	254	100.00	

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is yellow: A simple indicator to calculate. It provides a partial measure of science education in member states. It should be seen in relation to other indicators of science education more generally. It also relates to the other keys of RRI, because it is institutionalization of RRI in organisations. Specifying the question will provide variation that reflects more accurately the extent to which science education is emphasized and prioritized in organisations. If the indicator is deemed to be relevant for SUPER_MoRRI, then few changes can improve the indicator.

2.1.8   SLSE4: Citizen science activites in RPOs

Measurement and definition

SLSE4 captures if research performing organisations are engaged in citizen science in projects or through scientific publications about it.

1. Number of member organisations in the European Citizen Science Association (ECSA), and

2. The number of scientific publications concerning 'citizen science' (Retrieved from Scopus).

The indicator calculates the average of two variables: the number of memberships to ESCA per thousand researchers and the number of articles with the phrase "citizen science" in the abstract or title pr. 1000 researchers.

Congruency and relevance

The indicator provides a measure or the concept of citizen science and thus "co-production of knowledge". The indicator mainly shows the "successful" instances of co-production of scientific knowledge as it is based on published research. Other forms of knowledge that is not published, because it has other purposes (i.e. innovation or process improvement) or is unsuccessful is not included in the indicator.

Validity

The low prevalence of members in ESCA makes half of the indicator very susceptible to small changes in memberships. The second part approximates articles using citizen science and may exclude many that use citizen science methodologies but use other concepts to describe them.

Evaluation of validity, relevance and feasibility

The evaluation is Yellow: The relevance of the indicator is high as it provides a view to the degree to which citizen science is being adopted as a methodology in research performing organisations and member states. However, there are validity concerns when it comes to the pooling of two rather unrelated variables and the low frequency of especially memberships and the biases that can accompany them. A possible avenue forward is to focus purely on publications and extend and fine-tune the phrases used to identify articles that have employed a citizen science methodology.

2.1.9   PE5: Public engagement performance mechanisms at the level of research institutions

Measurement and definition

PE5 is based on Q26 and Q27 from the HEI and RPO survey; Q26 is a battery question with a total of 13 items and Q27 a question with three mutually exclusive responses.

Q26: Which of the following **mechanisms** does your institution apply in order to interact with citizens and societal stakeholders? Please consider whether there are changes in the practices of your institution over the years by providing answers for 2014, 2015, and 2016 (check those that apply) Research projects in partnership with non-academic organizations

1) Collaboration with NGO's and local government bodies (collab)
2) Participation in EU projects/networks about Public Engagement (partic)
3) Community representatives in boards or committees (represent)
4) Specific activities involving schools children visiting the institution (visit)
5) Meetings / conferences addressed primarily to the public (meeting)
6) Implementation of specific action plans targeting Public Engagement at your institution (implement)
7) Salary incentives for public outreach activities (salary)
8) Awards for science communication (award)
9) Availability of a press and/or Public Relations office (avail)
10) Public Engagement as a criterion for promotion (promote)
11) Public availability of information regarding completed and ongoing research activities (public_info)
12) Publications addressed primarily to the public (public_ps)
13) Organisation of outreach incentives such as 'open days' 'university festivals' etc (initia).

Q27: Which of the following statements come closest to the situation at your research institution? Please consider whether the priorities changed over the years by providing answers for 2014, 2015, and 2016

Public Engagement has high strategic priority at our research institution (3 pt.)
Public Engagement has moderate strategic priority at our research institution   (2 pt.)
Public Engagement is not a strategic priority at our research institution   (1 pt.)

Each confirmatory answer to the battery of questions in Q26 is coded as a 1 and each negative answer is a zero. The responses to Q 27 were given a value ranging from 1-3. The indicator is then calculated by

summing the points and normalizing to 0-1. The highest possible number of points before normalization is 16. PE5 is the normalized average score for organisations in the country.

Congruency and relevance

The questions relate to mechanisms the organisations use in relation to public communication, activism, deliberation, science education and other concepts. In addition, the perceived strategic priority of public engagement is a part of the indicator.

Validity

The questions in and of themselves have relevance, however pooling them and assigning an equal value to each makes this an indicator difficult to interpret. Any comparison between countries or over time will cause uncertainty to what causes differences.

Empirically, some of the activities are performed by almost all organisations. Moreover, if an organisation performs one activity they are also likely to perform another. Finally, if an organisation performs many activities they are also likely to respond that PE has a high strategic priority in their organisation.

Since the battery of questions about activities and strategic priority are highly correlated empirically and theoretically, it would be an option to choose one or the other and not combine them in an additive indicator.

Table PE5.1 shows the mean responses to Q26.

If organisations engage in the most far-reaching activities, they are also very likely to have engaged in the less fare-reaching activates. Table PE5.2 shows the mean responses to Q26 given that they answer yes to having awards for science communication. Table PE5.3 shows the mean of Q25 if an organisation provides awards, promotion and salary on the basis of PE. These organisations engage in all other activities (with very few exceptions)

This indicates that there is a degree of pyramidal shape when it comes to implementation of PE actions in organisations. This empirical finding relates to a theoretical expectation that the general and simple acts are performed before the more invasive and specific actions.

Table PE5.1

Variable	Obs	Mean	Std. Dev.	Min	Max
partner	370	.8810811	.3241314	0	1
collab	370	.8891892	.3143228	0	1
partic	370	.6054054	.4894253	0	1
represent	370	.6432432	.4796908	0	1
visit	370	.7810811	.4140735	0	1
meeting	370	.8189189	.3856068	0	1
implement	370	.4513514	.4983015	0	1
salary	370	.1486486	.3562235	0	1
award	370	.3297297	.4707515	0	1
avail	370	.7864865	.410342	0	1
promote	370	.2864865	.4527316	0	1
public_info	370	.8054054	.3964244	0	1
public_ps	370	.6972973	.4600498	0	1
initia	370	.7945946	.4045446	0	1

Table PE5.2

Variable	Obs	Mean	Std. Dev.	Min	Max
partner	122	.9836066	.1275067	0	1
collab	122	.9836066	.1275067	0	1
partic	122	.8360656	.3717427	0	1
represent	122	.795082	.4053062	0	1
visit	122	.8852459	.3200393	0	1
meeting	122	.942623	.2335207	0	1
implement	122	.647541	.479706	0	1
salary	122	.2868852	.4541727	0	1
award	122	1	0	1	1
avail	122	.9180328	.275446	0	1
promote	122	.4672131	.5009813	0	1
public_info	122	.9098361	.2875976	0	1
public_ps	122	.795082	.4053062	0	1
initia	122	.9262295	.262475	0	1

Table PE5.3

Variable	Obs	Mean	Std. Dev.	Min	Max
partner	30	1	0	1	1
collab	30	1	0	1	1
partic	30	1	0	1	1
represent	30	.9666667	.1825742	0	1
visit	30	1	0	1	1
meeting	30	1	0	1	1
implement	30	.9666667	.1825742	0	1
salary	30	1	0	1	1
award	30	1	0	1	1
avail	30	1	0	1	1
promote	30	1	0	1	1
public_info	30	1	0	1	1
public_ps	30	.9333333	.2537081	0	1
initia	30	1	0	1	1

Evaluation of validity, relevance and feasibility

The indicator measures a range of sub-concepts and provides a number that represents an overall score of PE mechanisms and strategic priority of PE. The relevance is very high. However, the indicator could

be either improved in its current form or changed to a number of related indicators. Both would have the purpose of improving the transparency of the indicator and improving the ability of the indicator to represent a more specific part of PE, rather than all sub-concepts added together.

The summary evaluation of the indicator is yellow: The relevance is high and the information given by the single questions is highly valuable. Yet, the diverse nature of the questions that are grouped create a cloudy interpretation of the indicator. Going forward, a simplification of the indicator, or splitting the indicator into smaller parts could improve transparency and congruency with the sub-concepts of PE.

2.1.10   PE7: Embedment of PE activities in the funding structure of PROs

Measurement and definition

PE7 is based on Q21 and Q23 in the RFO survey.

> Q21: Some research funding organisations contribute to Public Engagement through their funding schemes. Please indicate, if any of the following activities have been supported by targeted funding schemes in your organisation (Please tick all relevant boxes)

With the possible responses
1) Projects / activities which are primarily about disseminating research to citizens or societal stakeholders
2) Research projects which involve citizens or societal stakeholders in research activities
3) Research projects on Public Engagement (where the contents of the research is about Public Engagement
4) No such activites are funded through targeted schemes

> Q23: Please indicate the extent to which your funding agency has engaged with citizens and societal actors when developing its funding strategies

With possible responses: To a very large extent; To a large extent; To some extent; To a small extent; To a very small or no extent

The indicator is calculated by the sum of yes answers to Q21 plus a value of 1-5 for the response to Q23.

Congruency and relevance

The indicator consists of two relevant sub-concepts of PE. The first question relates to whether funding organisations fund research that has a strong PE focus. The second questions relates to whether the funding organisation itself engages in a dialogue with societal actors in order to shape funding strategies.

The first can give some form of understanding of how funding organizations see themselves and act (as actors for a government, owners etc.) or as representatives of society. The second, can give some indication of the degree to which research projects that emphasize different forms of PE are politically (within a country) emphasized as important areas.

Validity

When examining the empirical findings from the questions that make up the indicator, we see that there are four groups of funding organisations (see figure PE7.1): 1) No Public engagement in funding or in development of funding strategies, 2) Public engagement in funding and in development of funding

strategies, 3) public engagement in funding but not in development in funding strategies 4) no public engagement in funding but PE in development of funding strategies.

While the two first groups are the largest, the two others are not small. This means that when combining the two questions as an additive indicator, the two groups, 3 and 4 will receive scores in the same area, even though they are very different. While group 1 and 2 will be on each side of the spectrum. Thus by creating a composite indicator we lose information on what the "middle" consists of. Are they merely average on both concepts or high on one and low on the other?

Examining the questions closer, Q21 consists of three questions, one of which is whether the funding agencies fund projects about PE – few organisations answer that they do, and this may be because projects on PE are within specific fields that not all organisations fund. Therefore, it is worth considering whether this question is too specific.

Figure PE7.1

	Little or No Focus on PE for funding	Moderate to high Funding of Public engagement activities in research
Does not engage with societal actors	No PE Focus (24)(25)	Un-engaged with surrounding society – but funds PE activities in science (9)-(8)
Engages with societal actors	Engaged organisation, with no focus on engagement in science (15)-(16)	Engaged and funds engagement activities in science (26)-(25)

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is yellow: The indicator provides relevant illumination of PE activities within the funding environment. However, the indicator combines two parameters in a composite removing the ability to interpret the two aspects in isolation. The indicator could be divided into two "simple" indicators based on each of the questions, the two are two sides of the same concept but not related enough to warrant a composite indicator.

2.1.11   PE8: PE elements as evaluative criteria in research proposal evaluations

Measurement and definition

PE8 is based on Q24 in the RFO survey.

Q24: Please indicate the extent to which Public Engagement has been a criterion for the appraisal of research applications.

The possible responses were: To a very large extent; To a large extent; To some extent; To a small extent; To a very small or no extent
The response categories are coded from 5 to 1 (See distribution in table PE8.1).

Table PE.8.1

```
. tab q24_16

     q24_16 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         41       56.16       56.16
          2 |         11       15.07       71.23
          3 |         11       15.07       86.30
          4 |          6        8.22       94.52
          5 |          4        5.48      100.00
------------+-----------------------------------
      Total |         73      100.00
```

Congruency and relevance

The indicator asks generally about the use of PE in assessments. Therefore, the indicator reflects the degree of institutionalization or degree of attention to PE. In this sense, the indicator is congruent with the overall concept of PE.

It is a very relevant question that allows us to analyse the degree to which funding systems actively encourage PE principles though funding mandates in the countries.

Validity

The formulation of the question is very specific and can be easily understood and answered. The use of a Likert scale inevitably leads to the question of whether the values assigned are relevant. One option is to create a simple dummy variable that is one for funding agencies with large to very large extent and zero otherwise. This should be considered in relation to whether a change from no extent to a small extent can be considered to be equal to moving from some extent to a large extent.

From the MoRRI report the conclusion was that a "Reduction to a binary variable would reduce variation considerably". From the empirical data, we can see that this is mainly because very few RFOs indicate that they use PE in assessments to some, high or very high degree.

Similar to other questions in the survey, the question could instead read whether it is a standard criterion in all grants or a specific criterion in selected grants, or not a criterion in any grants. This could even be made as a Likert scale: PE criteria in all grants, in most grants, in some grants, in few grants, in very few or no grants.

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is yellow. It would be worth considering to create a new indicator that attempts to make the question relative, thereby creating a benchmark the respondent can answer high or low degree to. A possibility is to list a number of appraisal criteria and ask the respondents to rank or assign percentages in terms of how much weight they put on each criterion relative to each other. For example, relative to "research quality of applicants" to what degree does "planned public engagement activities" factor into the appraisal.

Alternatively, it could be worth considering to specify more clearly the ways in which PE can be a part of appraisal and ask whether these are used in the RFO.

The confusing empirical finding is that a large proportion indicate that they fund projects specifically aimed at PE (see PE7), but simultaneously indicate that PE is not an important appraisal criterion.

2.1.12   PE9: R&I democratization index

Measurement and definition

PE9 is based on Q5 and Q6 in the Science in Society actor survey. Q5 is a battery of five Likert scale questions (Strongly disagree to strongly agree).

> Q5.1) Citizens and civil society organizations are informed about developments in research and innovation
> Q5.2) Citizens and civil society organizations are consulted when political decisions about research and innovation are being made
> Q5.3) The opinions and advice of citizens and civil society organizations have a significant impact on political decisions about research and innovation
> Q5.4) The values and expectations of citizens and civil society organizations play an important role in setting the agenda for research and innovation
> Q5.5) My own organization has been able to influence decisions about research and innovation in my country.

Q6 is a binary question with answers "yes" and "no".

> 6) Are you aware of legal frameworks in your country which require participation of citizens and civil society organizations in science and technology decision making?

The calculation of the composite indicator is the normalized sum of weighted responses to the two questions. Q5 (1-5) and Q6 (0.1)

Congruency and relevance

Q5 asks about at least three different sub concepts related to PE. Q6 indicates whether the respondents are knowledgeable about whether or not there are national legal frameworks.
Perception of how inclusive and open the research and innovation system is, is a good indicator for Public engagement environment. This can be used to compare between countries and over time. Q6 provides an indication of a potential indicator. Since legal frameworks can either be or not, any variation within a country is because of the knowledgeability of the respondent. The relevance of the question is high, however, the validity of using respondents awareness of legal frameworks as proxies hereof can be discussed.

Validity

The first four questions are about the system in general while the last question is about the organisation the respondent represents. In general, the evaluation of the system (Agenda, Impact, Consulted and Informed) is more "negative" than the evaluation of the organisations' ability to influence the political process (Influence) (see tables below).

Table PE9.1

agenda	Freq.	Percent	Cum.
1	35	10.17	10.17
2	130	37.79	47.97
3	80	23.26	71.22
4	84	24.42	95.64
5	15	4.36	100.00
Total	344	100.00	

impact	Freq.	Percent	Cu
1	63	18.31	18.
2	137	39.83	58.
3	82	23.84	81.
4	52	15.12	97.
5	10	2.91	100.
Total	344	100.00	

consulted	Freq.	Percent	Cum.
1	55	15.85	15.85
2	147	42.36	58.21
3	58	16.71	74.93
4	80	23.05	97.98
5	7	2.02	100.00
Total	347	100.00	

informed	Freq.	Percent	Cum.
1	17	4.90	4.90
2	113	32.56	37.46
3	76	21.90	59.37
4	120	34.58	93.95
5	21	6.05	100.00
Total	347	100.00	

influence	Freq.	Percent	Cum.
1	21	6.05	6.05
2	46	13.26	19.31
3	79	22.77	42.07
4	142	40.92	83.00
5	59	17.00	100.00
Total	347	100.00	

To exemplify, only 18 % of the respondents agree or strongly agree that SiS organisations have significant impact on political decisions, while 57 % agree of strongly agree that their own organisation has been able to influence decisions about research and innovation.

The responses to question six exemplify the complexity of asking SiS actors about whether they are aware of any legal frameworks (see table below). Only in three countries (with 1 and 2 respondents) are the respondents in agreement on whether there exists or does not exits a legal framework. And in all countries except the three at least one respondents indicates that there is a legal framework (see table below).

Table PE9.2

1. Country	6. Are you aware of legal frameworks in your country which require participation		Total
	No	Yes	
Austria	17	7	24
Belgium	9	0	9
Bulgaria	9	3	12
Croatia	11	15	26
Cyprus	10	2	12
Czech Republic	21	5	26
Denmark	10	8	18
Estonia	9	7	16
Finland	5	5	10
France	5	7	12
Germany	11	10	21
Greece	3	5	8
Hungary	7	13	20
Ireland	2	0	2
Italy	6	3	9
Latvia	5	2	7
Lithuania	9	10	19
Luxembourg	1	0	1
Malta	7	1	8
Poland	1	3	4
Portugal	10	5	15
Romania	3	3	6
Slovakia	6	3	9
Slovenia	6	4	10
Spain	7	4	11
Sweden	4	5	9
The Netherlands	12	7	19
United Kingdom	3	6	9
Total	209	143	352

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is yellow: The indicator explores the degree to which SIS actors experience the ability and possibility of influencing, participating and being involved in decision making within R&I. Therefore, it is highly relevant. Creating an indicator on the basis of Q5 provides a good, comparable and robust indicator of the R&I environment in terms of openness and inclusiveness. Removing Q6 will improve the validity of the indicator.

While the indicator is important and relevant, it should be evaluated whether a SiS survey for the sake of two indicators (PE9 and PE10) is worth the time and effort. If so, it could be considered whether there is additional information that may be valuable to collect.

2.1.13   PE10: National infrastructure for involvement of citizens and societal actors in R&I

Measurement and definition

PE10 is based on Q8 (which is itself four questions) in the SiS survey.

Q8: Based on your experience and knowledge of the current situation in your country, please indicate the extent to which you agree with the following statements. The statement are as follows.

1) Citizens and civil society organisations have easy access to decision makers in the area of research and innovation policy
2) Citizens and civil society organisations are often represented in advisory bodies related to research and innovation policy
3) In my country, there are multiple channels for interaction between science and broader society
4) My own organisation plays an important role in mediating between science and broader society in my country

The responses to each question is coded from 1-5.

Congruency and relevance

The indicator is very relevant to understand and compare the national systems in terms of PE. The composite indicator does not inform about the single sub-concepts directly, rather they inform of the degree to which the respondents experience that national infrastructure facilitates public consultation, deliberation and participation at an aggregated level.

Validity

The main challenge is that the indicator is a composite and hides detailed information that is interesting and relevant. Q1 informs on the status of public consultation in a country, Q2 on state of public participation, Q3 informs on the status of infrastructure for participation, deliberation, consultation and communication while Q4, informs about perceived possibility for public participation and consultation.

Table PE10

```
      access |      Freq.      Percent        Cum.
-------------+-----------------------------------
           1 |         45        13.98       13.98
           2 |        113        35.09       49.07
           3 |         82        25.47       74.53
           4 |         73        22.67       97.20
           5 |          9         2.80      100.00
-------------+-----------------------------------
       Total |        322       100.00
```

. tab repres

```
representat |
        ion |      Freq.      Percent        Cum.
-------------+-----------------------------------
           1 |         39        12.19       12.19
           2 |        108        33.75       45.94
           3 |         83        25.94       71.88
           4 |         80        25.00       96.88
           5 |         10         3.13      100.00
-------------+-----------------------------------
       Total |        320       100.00
```

. tab interac

```
interaction |      Freq.      Percent        Cum.
-------------+-----------------------------------
           1 |         19         5.92        5.92
           2 |         79        24.61       30.53
           3 |         70        21.81       52.34
           4 |        127        39.56       91.90
           5 |         26         8.10      100.00
-------------+-----------------------------------
       Total |        321       100.00
```

. tab media

```
   mediation |      Freq.      Percent        Cum.
-------------+-----------------------------------
           1 |         13         4.06        4.06
           2 |         37        11.56       15.63
           3 |         63        19.69       35.31
           4 |        120        37.50       72.81
           5 |         87        27.19      100.00
-------------+-----------------------------------
       Total |        320       100.00
```

Evaluation of validity, relevance and feasibility

The evaluation is yellow: The indicator provides information on the "other side of the coin" of PE9. It is about the power and presence of channels for influence. It is so closely linked to PE9, that they could be combined or at least reported together.

2.1.14 E1.a: Ethics at the level of Universities and Public Research Organisations

Measurement and definition

The indicator is based on two questions in the RPO & HEI survey.

[does] your organisation have a Research Ethics Committee

[does] your institution have a Research Integrity Office

The indicator spans from zero to two. Zero is assigned if the answer is no to both questions, one if yes to one question and two if yes to both.

Congruency and relevance

It indicates the institutionalization of Ethical governance – because it measures the institutionalization of ethical discussions and deliberations of research procedures at the level of research performing organizations. The question is whether there is a formal infrastructure for dealing with ethics and research integrity.

It provides a good and informative way of illustrating the degree of institutionalization of practices within ethical research and Innovation.

Validity

The issue with the technical validity, is both the fact that there are two questions, which may change over time. A change of zero can actually mean that an organization removes the research integrity office and starts up and research ethics committee in the same period.

Second, the exact way in which an architecture is built up may be very different but still serve the same functions, therefore, the question could be broader: The organisation does not have a research ethics committee but has procedures for ethics, or the activities of a research integrity office are undertaken by another branch of the organization. One could argue that the indicator is dependent on the size of the organisation, even though small organizations have formal procedures they do not have stand-alone committees and offices.

Empirically, the findings show that very few organisations have Integrity offices, and if they do, they often also have an ethics committee. One could argue that since an integrity office is rare relative to an ethics committee, it should have a higher weight. Nevertheless, the indicator does not allow the analytical separation of two of the groups.

Table E1.a

```
                      type
         E1        HEI        PRO         Total
         ─────────────────────────────────────
          a          15          9           24
          b          51         90          141
          c          92         57          149
          d          55         13           68
         ─────────────────────────────────────
      Total         213        169          382
```

A= integrity office and no Ethics committee, B = No Integrity office and no ethics committee, C = no integrity office but ethics committee, D = Both integrity office and ethics committee

Evaluation of validity, relevance and feasibility

The evaluation is green: A good and simple indicator that can be relatively easily reported by the respondent. The validity concerns are minor and do not warrant changes to the indicator.

There is an additional indicator (E1.2). It attempts to formulate a more complex measurement of ethics based on sub-questions given only to organisations that have either an ethics committee or integrity office. The calculation of the indicator is not described in enough detail in order to evaluate.

2.1.15   E3.2: Research Funding Organisations Index

Measurement and definition

The indicator is based on Q25 in the RFO Survey.

> Has your organization integrated any types of ethics assessment/review in its funding decisions?

A yes response has a value of one and a no response has a value of zero.

Congruency and relevance

The indicator relates to the sub-concept of Ethical governance and reflects the degree to which the funding system has institutionalized ethical practices.

Validity

The only issue with the indicator is that the question in relatively vague. "Any types" opens up for a broad interpretation. Both funders with strict ethical assessment procedures and those who mention it in a paragraph can confidently answer yes. However, even though the phrasing of the question is reasonably broad only a little more than half respond that they do.

Table E3.2

```
25.3. 2016        Freq.      Percent        Cum.

        No            37        43.02        43.02
        Yes           49        56.98       100.00

      Total           86       100.00
```

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is green: The question is very relevant and straightforward. However, it may be possible to expand on the question by asking specifically, which types of ethical assessment are used. If such categories can be designed beforehand. This would allow the indicator to become more continuous in nature rather than binary. The reasoning behind this is that the response, "yes", can be closer or further from no, and grading the response would show this variation more clearly.

There is an additional indicator (E3.2). It attempts to formulate a more complex measurement of ethics. The calculation of the indicator is not explained in enough detail in order to evaluate.

2.1.16   OA1.1: Share of open access publishing

Measurement and definition

OA1.1 is based on publication data retrived form Web of Science merged with data on open access from Crossref, DOAJ and ROAD.

The indicator calculates the share of publications that are green or gold OA.

Congruency and relevance

The indicator provides a good indication of the degree to which open access is enacted in the member states.

Validity

The exact methodology of the indicator is not described in enough detail to evaluate the validity issues. However, given the data sources have a rather high coverage and the number of publications in each country is quite large there are few validity issues.

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is Green: A relevant and feasible indicator of open access. It should be examined whether this data is already compiled outside the scope of MoRRI and SUPER_MoRRI.

2.1.17   OA1.2 Share of gold open access publishing

Measurement and definition

OA1.1 is based on publication data retrieved form Web of Science merged with data on open access from Crossref, DOAJ and ROAD. The indicator calculates the share of publications that are gold OA.

Congruency and relevance

The indicator provides a good indication of the degree to which open access is enacted in the member states.

Validity

The exact methodology of the indicator is not described in enough detail to evaluate the validity issues. However, given the data sources have a rather high coverage and the size of publications in each country is quite large there are few validity issues.

Evaluation of validity, relevance and feasibility

The evaluation is Green: A relevant and feasible indicator of open access. It should be examined whether this data is not already complied outside the scope of MoRRI and SUPER_MoRRI.
It can be debated whether it is important to distinguish between green and gold open access for the purpose of MoRRI and SUPER_MoRRI.

2.1.18   OA3: Social media outreach/take up of Open Access Literature and open research data

Measurement and definition

Data collection has been based on publications from the Web of Science containing a DOI. DOIs have been matched with Altmetric.com and tweets and Wikipedia mentions have been extracted from this source. Two indicators were included:

      (1) The ratio of OA and non-OA publications mentioned through twitter per country;
      (2) The share of OA and non-OA publications cited in Wikipedia

Congruency and relevance

The indicator is interesting, however, the indicator is more informative for a discussion of what effects and benefits open access can have, rather than inform about developments in Open access activities and governance. Therefore, the use of this indicator depends on whether it is deemed as relevant to illuminate RRI.

Evaluation of validity, relevance and feasibility

The summary evaluation of the indicator is Yellow: The indicator is based on robust databases and data tools. The relevance as an indicator of RRI activities is however an open question.

2.1.19 OA6: RPO support structures for researchers as regards incentives and barriers for data sharing

Measurement and definition

OA6 is a composite indicator based on three batteries of questions in the RPO survey (Q51, 52 and 53) that have a total of 8 items. It is calculated by the average number of selected items by respondents in each country. It is normalized to a range between 0-1.

Q51: "Which of the following policies apply to your institution?
- Your institution has explicit open data management regulations
- Your institution has explicit institutional Gold or green Open access publishing regulations
- Your institution choses to follow funder or field specific incentives for open data and publication sharing

Q52: Which of the following open data sharing practices apply to your institution?
- Repositories are provided by your institution
- Repositories are provided by departments

Q53: Which of the following support options with regard to open access publishing and data sharing apply?
- The Library of your institution takes care of open access publishing
- Your institution provides IT support for FAIR data practices
- Your institution has specific budget for Open Access publishing
- Your institution has specific budget for the implementation of Open Data sharing
- Your institution provides support for on line communication (e.g. project websites) on publication and data sharing practices
- Your institution provides training in research data sharing e.g. about curation, metadata

Table OA6.1 Shows the indicator distribution for all organisations. Table OA6.2 shows the proportion of organisations answering yes to each question.

Table OA6.1

oa6	Freq.	Percent	Cum.
0	5	2.07	2.07
1	32	13.28	15.35
2	40	16.60	31.95
3	46	19.09	51.04
4	37	15.35	66.39
5	23	9.54	75.93
6	13	5.39	81.33
7	17	7.05	88.38
8	10	4.15	92.53
9	7	2.90	95.44
10	6	2.49	97.93
11	5	2.07	100.00
Total	241	100.00	

Table OA6.2

Variable	Obs	Mean	Std. Dev.	Min	Max
oa6	241	4.016598	2.580321	0	11
Q50_1	241	.3112033	.4639493	0	1
Q50_2	241	.2821577	.4509863	0	1
Q50_3	241	.3526971	.4788038	0	1
Q51_1	241	.6016598	.4905751	0	1
Q51_2	241	.1908714	.3938058	0	1
Q52_1	241	.560166	.4973999	0	1
Q52_2	241	.2946058	.4568141	0	1
Q52_3	241	.3278008	.4703887	0	1
Q52_4	241	.1742739	.3801343	0	1
Q52_5	241	.6182573	.486825	0	1
Q52_6	241	.3029046	.4604706	0	1

Congruency and relevance

OA6 is congruent with the overall concepts of open access and open data. The indicator provides a numerical approximation to the degree to which organisations have implemented support structures for Open Access and Open data.

Validity

The MoRRi report reported a Cronbach's alpha of 0.73 indicating a consistent composite. However, the intraclass test showed that there was more variation within than between countries implying that it is not an optimal country level indicator.

The additive nature of the indicator creates some quirks. For example, most organisations have data repositories either at the institutional level or the department level. However, some indicate that they have both. If they have both, the indicator is higher than if they only have one or the other. It can be debated whether having two systems should be counted as having "more" support structures. Second, it can be debated whether support structures for open access are meaningfully different between an institution that follows or does not follow funder or field specific incentives.

Evaluation of validity, relevance and feasibility

Evaluation is yellow: It is congruent with the overall concepts of open access and open data. The indicator provides a numerical approximation to the degree to which organisations have implemented support structures for Open Access and Open data.

In general, the indicator is complicated by the many questions, and could be simplified by reducing the number of items to the most central that are related to OA and OD, and that can be safely assumed to say something about the extent of support structures for open access and open data.

2.1.20 GOV2: RRI-related governance mechanisms within research funding and research performing organisations

Measurement and definition

GOV2 is a composite indicator that is based on a battery question posed in both the RPO and RFO survey.

> Q7: Based on your experience and knowledge, has your organisation established processes for managing the following aspects in 2014, 2015, 2016?"

Possible responses: Ethics; Citizen Engagement; Open Access; Gender Equality; Responsible R&I. Each organisation is assigned a value between 1 and 5 depending on the number of confirmatory responses.

Congruency and relevance

The indicator attempts to measure the degree of formal governance of RRI. A high value indicates that organisations have attempted to install processes to manage RRI.

Validity

Putting research organisations and funding organisations together is problematic, first because the relative number of responding RFOs to RPOs is different in each country. Thus, some variation may be because of this empirical fact. Most RFOs and RPOs answer yes to the questions. However, there is some variation in how many of the different key areas they respond to being active. In general RFOs are less likely to respond yes than RPOs.

Table GOV2.a - Results for RPOs

```
. sum ETH_16_1 PE_16_1 OA_16_1 GE_16_1 RRI_16_1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ETH_16_1	436	.6880734	.4638122	0	1
PE_16_1	436	.5527523	.4977806	0	1
OA_16_1	436	.6743119	.4691697	0	1
GE_16_1	436	.6513761	.4770821	0	1
RRI_16_1	436	.5986239	.4907399	0	1

Table GOV2.b - Results for RFOs

```
    Variable  │      Obs       Mean    Std. Dev.       Min        Max
──────────────┼─────────────────────────────────────────────────────
    ETH_16_1  │      105    .647619    .4800031          0          1
     PE_16_1  │      105   .5047619    .5023753          0          1
     OA_16_1  │      105   .5809524    .4957696          0          1
     GE_16_1  │      105   .4857143     .502193          0          1
    RRI_16_1  │      105   .4761905    .5018282          0          1
```

Evaluation of validity, relevance and feasibility

The indicator provides some indication of the degree to which organisations emphasise RRI and therefore, it can be used as an aggregate measure of governance of RRI. It is weak in terms of specificity, because the content and extent of activities within the keys is not specified.

2.1.21 GOV3: RRI-related governance mechanisms within research funding and performing organisations – composite index

Measurement and definition

The indicator is a composite index measuring the share of higher education institutions (HEIs/RPO) and research funding organizations (RFO) actively promoting or encouraging responsible research and innovation. The indicator is based on one question included in both the MoRRI RPO and RFO surveys. For both surveys, the question was: "Did your organization actively encourage the following among researchers, employees or partner organizations during 2016" RPOs and RFO were then presented with the following five examples:

Example 1A: Ethics in research and innovation.
Example 1B: Citizen engagement and participation of social actors.
Example 1C: Open access and open science.
Example 1D: Gender equality in research and innovation.
Example 1E: Responsible research and innovation.

The indicator was created by awarding each HEI or RFO the corresponding number of points per example. For each country, the individual scores were aggregated by computing the average, weighted by the response rate in each country, for HEIs and RPOs respectively.

Congruency and relevance

The indicator relates to the concept of governance by asking relevant organisations whether they have put governance mechanism in place for each of the five RRI Keys. The basis of the indicator is a relatively superficial question. However, it provides with an overall estimation of how much organisations have focused their attention and efforts to RRI. The relevance of the indicator at the country level can be questioned; yet at the organisational level, it provides a possibility to evaluate the degree of RRI governance mechanisms.

Validity

The main challenge for this indicator is the broadness of the questions. The interpretation of what constitutes "actively encouraging" can be understood differently for each respondent. It is a trade-off as providing specific examples will narrow the number of positive answers even though an organisation

does "activity encourage" RRI through governance mechanisms but not is the any of the examples provided.

Combining RFO and RPO survey responses also may hide interesting variation if the one type of organisation tends to be less likely to encourage RRI.

Evaluation

The summary evaluation of the indicator is Yellow: A simple indicator that allows for broad-based interpretation of RRI governance. Lacks specificity that helps to interpret the indicator.

2.2   Evaluation of indicators (secondary data)

2.2.1   GE2: Share of Women researchers by sector

Description

GE2 is the share of female researchers in the work force. The indicator is available for all sectors (GE2.1), the Business Sector (GE2.2), Higher Education sector (GE2.3) and the Government sector (GE2.4). The indicators show the development of women researcher participation in each sector relative to men researchers.

The indicator is extracted from Eurostat statistics on research and development (rd_p_femres).

Concept

The indicator relates to the concept of Gender Equality in R&I. The indicator provides an aggregate measure of how the participation of women researchers is developing over time in the member states and in different sectors.

Feasibility and possibilities

The data is available for all countries in all years, with some exceptions. Currently, most countries are covered in 2015, fewer in 2016 and only a hand-full in 2017. Presentation of the data can use either the most recent comparable year for all countries (2015) or the latest possible year for all countries.

A cursory inspection of the data shows that year-to-year fluctuations of +/-2 percentage points is normal. Therefore, a discussion of temporal developments should be performed with a long time line with temporal fluctuations in mind. The data goes back to 2008.

Evaluation

The summary evaluation is green: The indicator is relevant for tracking the overall developments in representation and participation of women researchers in R&I. The ability to disaggregate to different sectors also allows an inspection of where in the research and innovation system gender equality is lagging and improving. The data are easily downloaded from the Eurostat website and can be interpreted easily. Only minor issue are missing data in particular years for selected countries.

2.2.2 GE4: Dissimilarity index

Description

GE4 is an index calculated by She-Figures based on Eurostat data (rd_p_perssci) and is calculated every three years (latest year 2018).

The Dissimilarity Index provides a theoretical measurement of the percentage of women and men who would have to move to another field of science to ensure a gender balanced distribution across fields. It measures the distance from balanced gender distribution across fields for horizontal segregation in research

The indicator is available for both the government sector (GE4.1) and the Higher education sector (GE4.2)

Concept

The indicator represents the concept of gender equality in horizontal participation at the national level. The indicator provides a representation of how balanced representation is between scientific fields. Thus, it allows for an interpretation of the developments in stimulating equality within sub-fields of science.

Feasibility and possibilities

The data is available to download from She-Figures easily and demands little effort. The data is available for every third year.

Evaluation

The summary evaluation of the indicator is green: Overall, the indicator has a high level of relevance and validity. The data is easily extracted:

2.2.3 GE6: Glass ceiling index

Description

The Glass Ceiling Index measures women's chances of reaching the highest academic ranks relative to men's chances. It illustrates the difficulties women have to reach the highest organisational levels within RPOs. The proportion of women at academic levels A, B and C can be compared with the proportion of men at these levels. The share of women in Grade A as a comparison to the share of women in academia overall can be compared with the results for men. These data cover the higher education sector at the national level.

The Index is calculated by She-Figures every three years (latest year 2018).

Concept

The indicator represents the concept of vertical representation.

Feasibility and possibilities

The data is easily assessable through She-Figures website and covers every three years, with the latest year 2018.

Evaluation

The summary evaluation of the indicator is green: The indicator provides a possibility to track the developments in vertical representation in science. The data is available from the she-figures web site and is compiled every three years – latest available is 2018:

2.2.4   GE7: Gender Pay Gap

Description

The gender wage gap indicator measures gender variations with respect to annual and hourly earnings, and. It is calculated as the difference in gross annual earnings between women and men in relation to male gross annual earnings (interval)

The data is retrieved from Eurostat ISCO-08 code 2 and 3 – Academic Profession, Technicians and Associate Professionals

Concept

The indicator is used as a proxy for gender equality in the academic as well as the non-academic research sector. The wage differential is a product of both the different levels at which men and women are represented and potential biases in wage-setting and -negotiating.

Feasibility and possibilities

The data is easily retrievable from Eurostat.

Evaluation

The summary evaluation of the indicator is green: The indicator is a valid and relevant indicator of gender equality within R&I and is easily retrievable for all European countries over a long time period.

2.2.5   GE10: Share of female inventors and authors

Description

Data about publications (articles, letters, notes and reviews) and authors are extracted from Scopus (years 2005-16), data about transnational patents applications and inventors are extracted from PATSTAT (years 2005-15). Gender information is added by applying a gender identification method based on first names.

GE10 is presented as two indicators: GE10.1 Share of female authors and GE10.2 share of female inventors.

The number of patents in each country can be quite small for a given year. This can create large fluctuations in the indicator. Therefore, number of patents should be reported alongside the indicator. Alternatively, a longer time period than one year can be used.

Concept

The indicator represents the concept of gender equality in terms of innovative and scientific outputs. It is thus related to the notion of horizontal and vertical inclusion.

Feasibility and possibilities

There are similar measures available in She-figures. Therefore, it should be debated whether the investment in time and resources is worth it relative to simply using the she-figures data.

Evaluation

The overall evaluation of the indicator is yellow: Using She-Figures data will be easier and less costly. If specific questions are wanted then performing the analysis is also a possibility. The indicator relies on the ability to assign gender from first names. The precision of the algorithm should be reported alongside the indicator.

2.2.6   SLSE3: Science Communication Culture

Description

This composite indicator summarizes the overall national science communication culture. It was originally developed for the MASIS project. Data from the MASIS project, specifically the publication Mejlgaard et al (2012), Locating science in society across Europe: Clusters and conferences, Science and Public Policy 39, pp. 741-750 It builds on six parameters that collectively form a framework for describing the science communication culture of a specific country. These include, the degree of institutionalization (e.g. the presence of popular science magazines, regularity of science section in newspapers, dedicated science communication in television etc.), political attention to the field, the scale and diversity of actor involvement, traditions for popularization within academia, public interest in science and technology, and finally, the training and organizational characteristics of science journalism in the country.

Data collection is based on country reports produced by a network of national experts, following a common guideline and template.

Concept

The indicator relates to the concept of science communication. Specifically, the culture and national context surrounding science communication.

Feasibility and possibilities

In order to recollect data across countries, a setup similar to the MASIS project would be required. This involves national experts conducting desk research and interviews in their respective countries. The guidelines from the MASIS project could be adopted.

Evaluation

The summary evaluation of the indicator is red: While the indicator is both relevant and informative, the resources needed for the collection of a similar indicator is extremely high and the ability to compare over time is difficult because the indicator is a result of both data collection and qualitative data analysis.

The indicator can be used as inspiration for questions in a survey instead. Alternatively, country correspondents can be asked to perform a smaller data collection exercise to populate central elements of the indicator:

2.2.7 PE1: Models of public involvement in S&T decision making

Description

Models of public involvement in S&T decision making is a two-dimensional indicator. On one dimension is the degree of formalisation of structures and mechanisms, at the national level, for the involvement of citizens in decisions about science and technology. On the second dimension is the degree to which citizens are involved in making decisions. The two dimensions are considered to reflect the degree of overall democratization of science and technology decision-making. On the basis of these two dimensions, countries are grouped into a four-category typology.

The indicator is presented in Mejlgaard et al (2012): 'Locating Science in Society across Europe – Clusters and Consequences', in Science and Public Policy 39(6): 741-50, p. 746, table 3.

Feasibility and possibilities

The indicator is feasible as a one-off source. In order to recollect data across countries, a setup similar to the MASIS project would be required. This involves national experts conducting desk research and interviews in their respective countries. The guidelines from the MASIS project could be adopted.

Evaluation

Red: While the indicator is both relevant and informative the resources needed for the collection of a similar indicator is extremely high and the ability to compare over time is difficult because the indicator is a result of both data collection and qualitative data analysis.

The indicator can be used as inspiration for questions in a survey instead. Alternatively, country correspondents can be asked to perform a smaller data collection exercise to populate central elements of the indicator.

2.2.8 PE2: Policy-oriented engagement with science

Description

PE3 is an indicator based on responses to the Eurobarometer on RRI.
(http://data.europa.eu/euodp/en/data/dataset/S1096_79_2_401)

> 'What is the level of involvement citizens should have when it comes to decisions made about science and technology?'

Response categories: citizens do not need to be involved or informed; citizens should only be informed; citizens should be consulted and their opinions should be considered; citizens should participate and have an active role; citizens' opinions should be binding; and don't know.

The indicator reports the share of citizens at the national level expressing a preference for active participation. The indicator is calculated as a mean national score aggregated from a representative sample of citizens by country.

Concept

The indicator is related to the public's perception on public participation and deliberation while PE2 was focused actual participation. The indicator provides a view of the public opinion on how the public should be engaged in science and technology decision making.

Feasibility and possibilities

The indicator is feasible for application. However, continued future data collection would be expensive, unless aligned with the Eurobarometer series work

Evaluation

The summary evaluation of the indicator is yellow. Relevant and valid but depends on a new Eurobarometer.

2.2.9   PE3: Citizen preferences for active participation in S&T decision making

Description

This indicator is built as a composite measure based on three individual items from the 2010 Eurobarometer on biotechnology. It divides respondents into three categories depending on their responses to background items concerning GM food. The three categories of responses are:

1. "have heard and talked and/or searched for information";

2. "have heard but not talked or searched for information"; and

3. "have not heard".

The indicator taps into degrees of active information search, or what could be considered horizontal engagement, around controversial technologies.

Numerical value (share of citizens who have heard and talked and/or searched for information). The indicator is calculated as a mean national score aggregated from a representative sample of citizens by country.

Concept

The indicator relates both to public engagement and the actual level of engagement the public indicates to have, but also the indicator relates to science literacy.

Feasibility and possibilities

The indicator is feasible for application. However, continued future data collection would be expensive, unless aligned with the Eurobarometer series work

Evaluation

Green, a very relevant and valid indicator. The only issue is whether the question can be a part of the next Eurobarometer round.

2.2.10   OA4: Public perception of Open Access - PPOA

Description

The indicator on public perception of Open Access is constructed form a question in the Eurobarometer 2013. It provides the share of citizens in a country who think that publicly funded research should be made available to the general public.

Concept

The indicator relates to the concept of open access in terms of public opinion. Therefore, the indicator captures not the use or implementation of open access but rather the degree to which there is public support or demand for the implementation and use of open access.
The indicator thus relates to other RRI concepts such as science literacy and public engagement.

Feasibility and possibilities

The indicator is feasible for application. However, continued future data collection would be expensive, unless aligned with the Eurobarometer series work.

Evaluation

Yellow: The relevance of the indicator depends on how open access is viewed and the reasoning behind wanting to measure public perception of open access. The indicator is rather simple to construct, however demands that the Eurobarometer includes this question in a new wave.

2.2.11   OA5: Funder mandates

Description

The indicator presents if and how many funder mandates for open access publishing there are in the EU Member States. Funder/institutional mandates relate to the policy and practice of funding institutions giving research grants or of academic institutions to request the research output to be made openly accessible.
The source of the data and the calculation is not described in the available documents to an extent that enables an evaluation of the indicator. The source of the data is "openAire".

Concept

The indicator relates to the concept of open access and the degree to which open access is an institutionalized part of the funding process.

Feasibility and possibilities

The feasibility depends on whether a similar study will be performed and made available.

Evaluation

The summary evaluation of the indicator is red: The indicator has the possibility to provide a rough indication of the degree of open access implementation. However, the feasibility of recreating the indicator for comparisons over time are very questionable.

2.2.12 GOV1: Use of science in Policy making

Description

GOV1 is a nominal indicator developed in the MASIS project in 2012. It classifies the extent to which science is used in policy-making, with four different categories:

1 = No formalization and low impact

2 = Formalized and low impact

3 = Less formalized and considerable impact

4 = Highly formalized and high impact

Concept

The indicator relates to the national RRI governance landscape. It divided countries into four types in terms of how nations use science to inform policymaking.

Feasibility and possibilities

The indicator serves as a way to describe and discuss developments in countries in relation to RRI governance.

Evaluation:

The summary evaluation of the indicator is red: In order to recollect data across countries, a setup similar to the MASIS project would be required. This involves national experts conducting desk research and interviews in their respective countries. The guidelines from the MASIS project could be adopted.

2.2.13 E2: National Ethics Committees Index

Description

The indicator is based in data from a qualitative project "EPOCH"(https://epochconference2012.wordpress.com/about/). The index captures qualities of national ethics committee (NEC) infrastructure in a country. The index measures existence, output, impact and quality of NECs. It looks at the output in terms of opinions but also in terms of contributing to public debate, policy making. It particularly looks at the role of NECs by measuring the publication of work results, the organisation of public events, classification of existing public nvolvement mechanisms, involvement of target groups and the existence and quality of websites.

Index (from 0 to 1) constructed on the basis of set of qualitative criteria of the NEC. Final country score is the average score of all criteria:

- Publication of work results: "Always"=1; "Sometimes"=0
- Organization of public events: "Yes"=1; "No"=0
- Existence of specific public participation mechanisms: "Yes"=1;"No"=0
- Involvment of target groups: "Yes"=1; "No"=0
- Existence of websites: "Yes"=1; "No"=0

- Existence of well-organized websites providing information: "Yes"=1; "No"=0

Concept

The indicator provides a numerical ranking of the institutionalisation of ethics in a number of European countries.

Feasibility and possibilities

It would require using country correspondents to collect the data again.

Evaluation

The indicator only covers a handful of countries and the complexity is relatively high. Therfore, the evaluation is Yellow. A possibility is to use country correspodnents to collect the data. However, this would require significant work for the country correspondents.

2.3 Summary of indicator evaluation

Gender equality: In MoRRI a total of 15 indicators were compiled to measure the concept of gender equality. Five are original indicators created on the background of primary data, while ten are based on data provided by Eurostat and She-Figures. One indicator relates to the sub-concept of structural change in institutions (GE1), two indicators relate to the sub-concepts of gender content in research (GE3 & GE5) while twelve indicators relate to vertical and horizontal representation and participation in RRI.
In general, this RRI key is the most developed when it comes to measurement. The indicator-concept congruency is very high; each indicator is highly aligned with a sub concept within gender equality. This is also due to the substantial work being done within gender studies.

Looking forward, possible ways to reduce the number of indicators is to delimit the number of indicators related to horizontal and vertical representation and participation. Gender in research content and structural change are areas with a relative lack of indicators.

The indicators based on secondary data are generally of a sound level when it comes to validity and relevance. The indicators based on primary data generally have some minor weaknesses that need to be addressed in SUPER_MoRRI.

Science literacy and Science education: In MoRRI a total of four indicators were compiled to measure the concept of Science literacy and science education. Three are original indicators based on secondary data and country correspondents. One indicator is taken from the MASIS project. The four indicators each represent one of the three sub concepts of Science education (SLSE1, SLSE2), science communication (SLSE3) and co-production of knowledge (SLSE4).

Looking forward, the potential gaps are science communication where one indicator based on the MASIS project will be difficult to reproduce without serious investment. Therefore, new indicators should be considered. Reproducing the remaining three indicators would only demand minor improvements.

Public engagement: In MoRRI, a total of nine indicators were compiled to measure the concept of Public engagement. Five based on primary data and four on secondary data. The indicators are mainly at the general level of public engagement in terms of governance and activities and do not relate to single sub-

concepts. Exceptions include PE2 and PE4 which measure public participation and PE3 that measures citizens' attitudes toward public engagement.

Eight out of the nine indicators can be repeated with minor to major changes, while PE1, based on the MASIS project, would demand a large amount of resources not immediately available to the SUPER_MoRRI project. In addition, a number of indicators require that the Eurobarometer includes the same questions in a new round.

Open access: In MoRRI a total of six indicators were compiled to measure the concept of open access. The indicators reflect mainly open access publishing and institutionalized structures for open access publishing and data storage. None of the indicators measure open data as a stand-alone concept. Four of the indicators rely purely on secondary data to create the indicators, while one is based on primary data collection (OA6). Five out of six indicators are feasible for continued monitoring given minor changes and improvements.

Going forward, a focus on open access publishing (OA1) organisational governance of open access (OA6) and public perception of open access (O4) are likely to be the most feasible. An attempt to broaden the knowledge of open data should be attempted.

Ethics: In MoRRI a total of five indicators were compiled to measure the concept of Ethics. They focus mainly on the governance structures in organisations (E1 & E3) and nationally (E2). Two indicators are very complicated and are not sufficiently described to warrant an evaluation, however they attempt to measure the same as E1.1 and E3.1. In general, ethics is one of the most thinly covered RRI keys. Going forward, E1 and E3 are feasible to repeat, while work on developing E2 may be plausible. All three indicators can be considered to be rather shallow, as they focus on the existence of structures for ethical research.

Governance: In MoRRI a total of three indicators were compiled to measure governance of RRI. Two indicators are based on primary organisational data and one indicator is borrowed from the MaSiS project. The first two are feasible to reproduce while the MASIS indicator would require both a rethinking of data collection and a large amount of resources to repeat.

Going forward a discussion of how best to measure and view governance should be had, since governance is indirectly covered through a number of the indicators that are under the other keys.

In table S1 a simple summary of each indicator is provided. A green, yellow or red code is given overall and in terms of relevance, validity and feasibility.
- Green: The indicator can be collected without any major improvements or resources.
- Yellow: There are issues with relevance, validity or feasibility that needs to be addressed before collecting data.
- Red: There are significant constraints that make data collection and indicator construction very difficult or unfeasible.

Table S1.

Indicator	Sub-concept(s)	Source (P/S)	Lowest/highest level of aggregation	Evaluation	Relevance	Validity	Feasibility
GE1	Structural Change / Governance of GE	HEI/RPO (P)	ORG/COUNTRY	G	G	G	G
GE2.1	Vertical/Horizontal Integration in R&I	EUROS (S)	COUNTRY	G	G	G	G
GE2.2	Vertical/Horizontal Integration in R&I	EUROS (S)	COUNTRY	G	G	G	G
GE2.3	Vertical/Horizontal Integration in R&I	EUROS (S)	COUNTRY	G	G	G	G
GE2.4	Vertical/Horizontal Integration in R&I	EUROS (S)	COUNTRY	G	G	G	G
GE3	Gender in research content	RFO (P)	ORG/COUNTRY	Y	G	Y	G
GE4.1	Horizontal integration in R&I	SHEFIG (S)	COUNTRY	G	G	G	G
GE4.2	Horizontal integration in R&I	SHEFIG (S)	COUNTRY	G	G	G	G
GE5	Gender in research content	HEI/RPO (P)	ORG/COUNTRY	Y	G	R	G
GE6	Vertical Integration In R&I	SHEFIG (S)	COUNTRY	G	G	G	G
GE7	Vertical/Horizontal Integration in R&I	EUROS (S)	COUNTRY	G	Y	G	G
GE8	Vertical Integration In R&I	HEI/RPO (P)	ORG/COUNTRY	G	G	G	G
GE9	Vertical/Horizontal Integration in R&I	HEI/RPO (P)	ORG/COUNTRY	R	G	R	Y
GE10.1	Vertical/Horizontal Integration in R&I	SOPUS/ (S)	Paper/Country	G	G	G	G
GE10.2	Vertical/Horizontal Integration in R&I	PATSTAT (S)	Patent/Country	G	G	Y	G

SLSE1	Science Education	CC - DESK RES (P)	COUNTRY	Y	G	G	Y
SLSE2	Science Education	HEI/RPO (P)	ORG/COUNTRY	G	G	G	G
SLSE3	Science Communication	MASIS (S)	COUNTRY	R	G	Y	R
SLSE4	Co-production of knowledge/citizen science	SCOPUS / ESCA (P/S)	Paper/Country	Y	G	Y	Y
PE1	PE General	MASIS (S)	COUNTRY	R	G	Y	R
PE2	Public participation	EUROBA (S)	COUNTRY	G	G	G	Y
PE3	PE perception	EUROBA (S)	COUNTRY	Y	Y	G	Y
PE4	Public participation; Science literacy	EUROBA (S)	COUNTRY	G	G	G	Y
PE5	PE General	HEI/RPO (P)	ORG/COUNTRY	Y	G	Y	G
PE7	PE General	HEI/RPO (P)	ORG/COUNTRY	Y	G	Y	G
PE8	PE General	HEI/RPO (P)	ORG/COUNTRY	Y	G	Y	G
PE9	PE General	SiS (P)	ORG/COUNTRY	Y	G	Y	Y
PE10	PE General	SiS (P)	ORG/COUNTRY	Y	G	Y	G
E1.1	Ethical governance	HEI/RPO (P)	ORG/COUNTRY	G	G	G	G
E1.2	Ethical governance	HEI/RPO (P)	ORG/COUNTRY	R	R	R	R
E2.1	Ethical Governance	EPOCH (S)	COUNTRY	Y	G	Y	Y
E3.1	Ethical governance	HEI/RPO (P)	ORG/COUNTRY	G	G	G	G
E3.2	Ethical governance	HEI/RPO (P)	ORG/COUNTRY	R	R	R	R
OA1.1	Open access (use, implementation)	WOS;MEND; CROSSREF (S/P)	Paper/Country	G	G	G	G
OA1.2	Open access (use, implementation)	WOS;MEND; CROSSREF; OPENAIRE (S/P)	Paper/Country	G	G	G	G
OA2	---						

OA3	Open access (uptake)	WOS;ALTMETRIC (S/P)	Paper/ Country	Y	Y	Y	Y
OA4	Open Access (pub. Perception)	EUROBA (S)	COUNTRY	Y	G	G	Y
OA5	Open Access (institutionalization)	OPENAIRE	COUNTRY	R	G	Y	R
OA6	Open access & Open data (structures)	HEI/RPO (P)	ORG/COUNTRY	Y	G	Y	G
GOV1	RRI Governance	MASIS (S)	COUNTRY	R	G	Y	R
GOV2	RRI Governance	HEI/RPO (P)	ORG/COUNTRY	G	G	G	G
GOV3	RRI Governance	HEI/RPO (P)	ORG/COUNTRY	Y	G	Y	G

2.4   General and overarching issues and challenges

2.4.1   Data availability & collection feasibility

Some of the indicators in MoRRI are borrowed from qualitative projects such as MASIS. The development of these indicators took a significant period of time and many resources. In order to reproduce these indicators we need to either repeat the same work or find alternative methods to collect similar information.

In MoRRI She-Figures, Eurostat and Bibliometric databases were used extensively for indicator construction. These are cheap and valid sources of information that we should look to in SUPER_MoRRI.

Finally, three surveys and one Eurobarometer were used for many of the indicators. The SiS survey only accounts for two indicators. The costs and benefits of sending out surveys should be discussed. The RPO survey is the most important in terms of reproducing the indicators.

2.4.2 Response rates

Many of the national indicators in the MoRRI monitoring system are based on aggregated organisational data (RFO survey, SiS survey and RPO survey). This does not in and of itself pose any issues. However, in order to have stable and reliable national level indicators, the total number of organisational responses within a country is very important. The number of responding organisations in each country varies between 1 and 16. Calculating national indicators demands that there is a sufficient number of relatively representative group of organisations, in order to generalize to the total population and thereby make comparisons over time and between countries. In order to obtain valid and consistent indicators, the response rate is also very important. If the responses in each time-period are from different organisations the chance of getting different indicator values also increases. The higher the response rate the more reliably the indicators can be compared over time and between countries. As it is not possible to force responses from a representative sample of organisations in a country, relatively large numbers are needed to ensure that we can consistently compare indicators.

Alternatively, the indicators can be calculated and reported at the organisational level, which would solve the issue of attempting to generalize to the national level.

2.4.3   Indicator coverage of concepts and sub-concepts

The indicators for each key are very different in terms of how broadly/narrowly they measure sub-concepts. A discussion of whether an approach such as the MoRRI projects should be undertaken, or whether a more mechanical approach should be taken (i.e. one narrow and specific indicator for each sub-concept).

2.4.4 Time comparisons

An important part of SUPER_MoRRI is to continue the time series of important indicators. Changes to the collection and specification of indicators will affect an accurate time comparison, therefore this has to be taken into account both in the development of the monitoring system and in the reporting of the results.

2.5 Suggestions for future monitoring system

2.5.1   Reduction of amount of data collection exercises and indicators

Based on the work in this document there are a number of concepts that are populated with very similar indicators. At the same time other concepts such as Ethics, Governance and Open access are thinly populated. There is a possibility to thin out in the number of indicators and focusing on what is found to be most informative.

2.5.2   Simplifying and aligning indicators with single sub-concepts

A number of indicators are relatively complicated both in terms of calculation, but also in terms of interpretation. Therefore, an option is to disaggregate some of the indicators and report more simple indicators. This can ease interpretation

2.5.3   Employing a multi-level approach to indicator construction and reporting

In MoRRI, the indicators were all aggregated to the country level. However, in SUPER_MoRRI we have the opportunity to choose the level of aggregation that is most informative. An option would be to present national indicators at a broad level for each of the keys, and then present organisational and regional indicators to provide a more nuanced representation of RRI

2.5.4   Improve responsibility through reporting on uncertainty and variability

The indicators in MoRRI are mainly reported as averages and countries are ranked accordingly. In order to promote responsible use of indicators, it could be proposed that we provide uncertainty intervals (based on responses, response rates and variation) and information on low response rates in tables. This is especially important when aggregating organisational or individual data to the country level, where there is great variation within a country.

SUPER MoRRI

Scientific Understanding and Provision of an Enhanced and Robust Monitoring system for RRI **Horizon 2020, Science with and for Society Work Programme 2018-2020,** Topic: SwafS-21-2018 **Grant Agreement Number**: 824671