The authors of this chapter

"RRI can perhaps be considered as a new paradigm of responsibility that goes beyond the traditional emphasis on fault and punishment, risk and compensation, uncertainty and precaution."

2

RRI as a governance paradigm: What is new?

Simone Arnaldi, Guido Gorgoni, Elena Pariotti

This chapter frames Responsible Research and Innovation (RRI) as an emerging governance approach in the EU regulatory context. We argue that the reference to fundamental rights makes RRI a distinctive approach to responsibility compared to other existing paradigms and that human rights, in particular those laid down in the Charter of Fundamental Rights of the European Union, are not necessarily a constraint but can instead be a catalyst for innovation. We maintain that a governance framework based on the complementarity between legal norms and voluntary commitments might successfully combine the respect for fundamental rights with the openness and flexibility of the innovation process.

1.5 RRI and the governance of technology[1]

RRI deal with situations in which knowledge is uncertain and consent is contested, so that traditional approaches addressing responsibility *ex post facto* by the means of liability or compensation are unsatisfactory. Instead, RRI promotes a more comprehensive approach to responsibility.

Academic literature and public debates alike have increasingly acknowledged the pervasiveness of uncertainty in science, technology and their governance. Uncertainty is no longer viewed as a residual area of ignorance and risk to be gradually reduced by way of increasing expert knowledge and enhanced technological control. It is rather a consequence of the ecological nature of technology, which cannot be eliminated, and that its interaction with the environment generates (Luhmann 1993).

As a consequence of the difficulty to predict future developments and possible risks, we are often only able to learn about these developments after technologies have been introduced and have shown their consequences for society. This way we enter into the domain of "manufactured risk" (Giddens 1999) and the unavoidable "secondary consequences" of action (Beck 1999). Indeed, this increasingly manipulative knowledge of nature and society produces uncertainty rather than reduces it, and this radical uncertainty reshapes the boundaries between science and policy. Knowledge and technology, therefore, implicitly incorporate models, world views and societal patterns (Wynne 1995), so that "the ways in which we know and

1 All the authors outlined the structure of the chapter. S. Arnaldi wrote Section 2.1; G. Gorgoni wrote Section 2.2; E. Pariotti wrote Section 2.3; all authors wrote Section 2.4. The authors have read and approved the manuscript.

represent the world (both nature and society) are inseparable from the way in which we choose to live in it." (Jasanoff 2004: 2)

Today, the governance of new technologies is therefore designed and implemented in situations that are characterised by uncertain knowledge and embattled consent (Douglas and Wildavsky 1983). We argue that RRI can be an effective answer to this twofold uncertainty, so that responsiveness and the normative steering of research and innovation acquire more importance over risk individuation and management.

The nature of RRI as a conceptual and policy approach aimed at actors' reciprocal responsibilisation, defines a space for innovative forms of governance centred on the adoption and the practical implementation of (self-)regulatory instruments such as codes of conduct, guidelines, technical standards, reporting, and audits.

Broadly speaking, soft regulation is a set of explicit rules, which have either a non-binding character or are utterly voluntary (Fredriksson et al. 2011, Skjærseth et al. 2006). Soft norms have an acknowledged legal relevance, though they lack a formally binding effect, precision, and clearly top-down delineated enforcement mechanisms (Shaffer and Pollack 2009). Because of this characteristic nature, soft norms have often been defined as "non-legislative modes of policy-making" (Hérriet in Fredriksson et al. 2011: 53) or even as "quasi-legal instruments" (Koutalakis et al. 2010: 330). Soft regulation describes a shift "from direct intervention ("rowing") to indirect intervention ("steering") in terms of enabling, motivating and pressing the regulated parties to regulate and to comply with self-regulation" (Dorbeck-Jung and Shelley-Egan 2013: 56).

Soft regulation is useful to regulators as it constitutes a tool for leveraging the information advantages of those actors who need to be regulated. This is considered an important asset in emerging technological fields that are characterised by a high degree of uncertainty and for which regulators lack the resources or information needed to develop sound "discretion-limiting rules" of a mandatory nature as it is:

"particularly the case in highly technical areas where the state depends on individual producers for crucial regulatory information related to product characteristics and production processes." (Koutalakis et al. 2010: 334)

In this context, soft regulation is used in processes where "there is the need to build a participated consensus on legal and political decisions" (Pariotti 2011: 516) and the institutional and organizational configurations of regulatory actions:

"provide little space for different and conflicting interests to be articulated. This does not mean that conflicts disappear, but that they may take other routes, or are put 'on hold', as it were." (Garsten and Jacobsson 2013: 422)

The expansion of soft regulation does not replace hard law as such, but creates "hybrid" regulatory frameworks; this happens when, for instance, a voluntary good practice code is used as a benchmark for compliance with a "hard law" prescription" (Heyvaert 2009: 650) or, on the contrary, when hard law is referred to in broader soft regulatory instruments. We maintain that this complementarity is just the kind of result that is pursued by the idea of RRI.

Table 2–1: Soft regulatory initiatives: some examples (source: Arnaldi 2014)

		Level of initiative	
		National / subnational	International / supranational
Initiator	Public	Voluntary Reporting Scheme for Engineered Nanoscale Materials (UK) (DEFRA 2008a, 2008b) EPA Nanoscale Materials Stewardship Program (EPA n.d.)	OECD Working Party on Nanotechnology (n.d.) European Commission Code of Conduct (2008)
	Private	Responsible Nanocode (n.d.)	ISO TC 229 (ISO n.d.) ResponsibleCare (ICCA 2006)

1.6 RRI and the evolutions of responsibility

RRI has to be examined in the context of the diversity and historical evolution of the notion of responsibility. Indeed responsibility is "a syndrome of concepts" (Vincent 2011) variously interconnected (e.g. Davis 2012, van de Poel 2011, Vincent 2011, Gorgoni 2011, Ricoeur 2000, Hart 1968). The different meanings of responsibility can be referred to as two distinct poles: a passive pole, relating to the imputation of responsibility (being held responsible) and an active pole, which is that of the voluntary preventive assumption of responsibility.

Indeed, responsibility can be equally understood in terms of the obligation to bear the consequences of an action (*liability*), as the capacity to act taking into account one's duties and giving an account of them (*accountability*), or as the capacity to act without relying on general pre-established rules or waiting for ex-post accounts, but rather by taking into account the specific context (*responsiveness*).

The idea of responsiveness is different from that of *reaction* typically associated with responsibility and is closer to the idea of a *response*, therefore characterising the idea of responsibility as both *open* and *active*:

"Response entails previous listening to a question. It entails openness, a willingness to understand and confront the other's commitments and concerns with ours, to look for a possible terrain of sharing. It entails readiness to rethink our own problem definition, goals, strategies, and identity." (Pellizzoni 2004: 557)

The distinction between the active and the passive modalities of responsibility implies the distinction between the temporal directions of responsibility, namely the retrospective and the prospective (Cane 2002).

Retrospective responsibility, or "historic responsibility" (Bovens 1988), is backward-looking, i.e. past-oriented, and is essentially linked to the idea of a *reaction*, which shapes the idea of responsibility in terms of sanction, compensation or justification. Responsibility in this case is called "retrospective" in that its key moment is the *ex post* evaluation of a situation.

Prospective responsibility is forward-looking, i.e. future-oriented, and is essentially linked to the idea of *assuming* and *exercising* responsibility, certainly in the sense of complying with the duties associated with our roles, but also by (pro)actively assuming responsibilities when the contents of our duties and tasks are not (or cannot) be established in advance. Responsibility is called "prospective" in that responsibility is not an ex-post judgement over a certain state of affairs, but a *projection* onto it, i.e. with no judgement in terms of a subsequent fault or compensation, but rather in terms of commitment. This active understanding of responsibility is central in regulatory strategies based on *responsibilisation*, intended as "predisposing actors to assume responsibility for their action" (Dorbeck-Jung and Shelley-Egan 2013: 60).

Considering the two semantic poles we described above and the predominant time dimension the different understandings of responsibility refer to, different *paradigms of responsibility* can be distinguished, according to their changing logic in combination with these elements. In our view, three main paradigms can be identified, all of which coexist despite the fact they were developed under specific historical conditions and therefore they do indeed characterise some typical "eras" of responsibility. By revisiting the work of François Ewald, we distinguish between the following:

1. The paradigm of fault, corresponding to the traditional moral and legal idea of responsibility as linked to a faulty causation by the agent. This paradigm of responsibility is essentially retrospective as it is based on the ex post judgement of a past action, and possibly on its sanction, and characterises both the legal and the ethical field (e.g. Hart 1968).

2. The paradigm of risk, in which the focus of responsibility is put on guaranteeing victims against damages (without reference to anybody's fault), rather than on sanctioning the "responsible" person(s), whose involvement in producing or not the damage becomes irrelevant under the "objective" logic of compensation. This model of responsibility is indeed prospective in that it aims at anticipating the occurrence of damages by the means of risk management techniques (Beck 1999). This way responsibility is turned towards the

future disclosing opportunities for action (otherwise "tied" by the spectrum of fault); but at the same time it remains linked to a retrospective logic in that it anticipates the occurrence of damage but it does not imply a higher responsibilisation of the practices concerned, as responsibility is based on statistical and not on ethical or legal criteria. Thus, paradoxically a sort of de-responsibilisation in terms of commitment is induced.

3. The paradigm of safety, as a reaction to a situation of uncertainty that cannot be domesticated by means of risk calculation. This paradigm was inaugurated by the development of the idea of precaution, both in the ethical and in the legal sense. Indeed the two former paradigms of responsibility are seriously challenged by the evolution of science and innovation, as they both presuppose either an identifiable author (fault) or some reliable data on which calculations (risk) are based, whilst contemporary science is characterised by uncertainty, as the direct or indirect outcomes of innovation practices cannot be fully anticipated (e.g. the effects of the use of chemical products in agriculture and their effects on the ecosystem, the effects of GMO's on the biosphere, etc.). Therefore the preventive approach of risk management cannot provide acceptable answers, nor would the fault paradigm help in making innovation processes more responsible. Within this context of uncertainty the focus of responsibility is put on anticipating the undesirable outcomes of techno-scientific activities, basing responsibility on value-centred decisions in a context of uncertainty rather than on a risk-based approach.

Those paradigms of responsibility coexist, overlap and sometimes compete with each other. When compared to the RRI idea, it presents some distinctive features that we should briefly analyse.

Despite some differences, the literature on RRI shares a largely common understanding of responsibility and its dimensions (see von Schomberg 2013, Owen 2014, van den Hoven et al. 2013, Forsberg et al. 2015):

- Responsibility is oriented to the future: the specific approach of RRI does not aim only at sanctioning, compensating or preventing the negative consequences of innovation; it aims indeed at steering the innovation processes according to societal values and needs, therefore advocating a *prospective* idea of responsibility.

- Responsibility is more proactive than reactive: responsibility is intended to be mainly a driving factor of the innovation process rather than a constraint, therefore it goes beyond the boundaries of what is legally due and relies on proactive anticipatory interventions.

- Responsibility is a collective and participative process: rather than being individual, responsibility is shared across different actors with different roles and powers along the innovation process, engaging with the collective shaping of societally acceptable research and innovation trajectories.

- Responsibility is plural: RRI links different dimensions of responsibility, namely the political, legal, ethical, and economic. Indeed the pursuit of responsible innovation rests on the voluntary adoption of standards which are not legally binding (ethical dimension of responsibility). These standards may become the normative references for RRI activities (political dimension of responsibility), so that our current "grand challenges" can be answered (social dimension of responsibility) respecting and promoting EU Fundamental Rights (legal dimension of responsibility) at the same time.

These features seem to set RRI apart from the other responsibility paradigms we have briefly described above (see Table 2–2 for an unavoidably simplified comparison). It does not mean that it *replaces* the other ones; rather it combines some of their elements in an original, and more comprehensive, fashion. Indeed, RRI can perhaps be considered as a new paradigm of responsibility that goes beyond the traditional emphasis on fault and punishment, risk and compensation, uncertainty and precaution, as it aims at *steering the innovation process from the inside* towards societal goals rather than coping with its (actual or anticipated) unwanted and unintended externalities.

Table 2–2: RRI and the evolution of responsibility paradigms

Paradigm	Fault	Risk	Safety	RRI
Criterion of ascription	Liability	Damage	Uncertainty	Responsiveness
Mean of realisation	Sanction	Compensation	Precaution	Participation
Target	Negative outcomes	Negative outcomes	Negative outcomes	Negative and positive outcomes
Dimension	Individual	Systemic	Collective	Collaborative
Orientation in time	Retrospective	Prospective / Retrospective	Prospective / Anticipative	Prospective / Proactive
Regulating mechanism	Hard law	Hard law	Hard law / Soft law	Self-regulation / Soft law / Hard law

1.7 RRI as a governance paradigm

RRI aims at actors' reciprocal responsibilisation, opening to innovative forms of governance centred on the adoption and the practical implementation of (self-)regulatory instruments such as codes of conduct, guidelines, technical standards, reporting, and audits. These types of regulatory instruments and their incorporation into hybrid regulatory schemes promote participation and power sharing, the integration of different levels of governance, diversity and decentralization, expansion of the space for stakeholders' deliberation.

RRI comprehensively combine and integrate various earlier approaches and methods, as:

"technology assessment and foresight, application of the precautionary principle, normative / ethical principles to design technology, innovation governance and stakeholder involvement and public engagement [in both deliberation and regulation]." (von Schomberg 2013: 65)

The literature that is most close to the EU policy environment from which the notion of RRI originates, includes fundamental rights as the source of orientation of research and innovation (von Schomberg 2013, Ozolina et al. 2009). In its most cited definition, RRI is defined as:

"a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society)." (von Schomberg 2011: 9; 2013: 63)

In this view, ethical acceptability "refers to a mandatory compliance with the fundamental values of the EU Charter on fundamental rights" (von Schomberg 2013: 63). Moreover, social desirability "captures the relevant and more specific normative anchor points of the treaty on the European Union" (von Schomberg 2013: 64).

Competitiveness, scientific progress, fundamental rights, environmental protection are among the normative anchor points of EU research and innovation policies and, therefore, it seems reasonable that they play a role as the normative "building-blocks" of a governance framework.

The definition of RRI we have cited grants a role to the legal dimension of RRI, and, above all, emphasizes the integrated presence of multiple dimensions within the notion of RRI, like the ethical, political, social and legal ones. The reference to fundamental rights could be

regarded as a way to rigidly set values and goals, even regardless of the development of public debate and public opinion. From this point of view, fundamental rights could be considered as normative constraints defined in a top-down way, limiting the scope and influence of public involvement. However, this representation of fundamental rights in general and of their specific role in RRI is indeed debatable.

Human rights are usually seen from two opposite perspectives and both of them should be rejected. According to a first view, human rights are abstract ideals, which can easily be reduced to rhetorical appeals. According to a different one, human rights are expressed by norms concerning solely the relationships between citizens and their governments or judicial courts. In this understanding, fundamental rights have no relation to public opinion.

Fundamental rights can, on the contrary, be thought of as claims that are justified by strong moral reasons and supported by legal norms, suitable to regulate both the relations between the government and the citizens ("vertical dimension"), and those between private actors ("horizontal dimension").

However, it is important to note that the legal norms supporting such claims are often structurally vague, because they have to apply to as many cases as possible.

The content of those fundamental rights is not established once for all in the law-making process, but must be shaped, also in a bottom-up manner and by several relevant actors during the application stage, like judges but also private actors promoting tools of self-regulation.

In science, technology and innovation, many private actors actively self-regulate as they possess the relevant information and knowledge, so that the contents of fundamental rights should emerge in a bottom-up fashion. Therefore, it is possible to maintain that fundamental rights are a basic reference for RRI and that, nevertheless, the development and implementation of such a model should and can come to terms with different values and with different interpretations of the rights themselves.

When understood in this way, human rights gain a central space in RRI as they affect the regulation of innovation in several ways:

1. on a judicial level: the judicial stance contributes to the definition of the content of rights.
2. on a policy level: the protection and promotion of rights act as a driver for policy making.
3. the reference to human rights plays a role also on a horizontal level, between private actors, like, for instance, when the most diverse organizations adopt and implement social responsibility instruments (codes of conduct, self-regulations).

Considering fundamental rights as essential elements of RRI does not imply the narrowing of the scope and of the role of public involvement in defining the objectives of research and innovation and their social acceptability. It does not mean that the normative standards to be complied with and the goals to be pursued are already fully set in a top-down manner. Far from it, once listed, fundamental rights have to be filled with contents and have to be detailed with regard to specific domains, contexts, and cases by the means of an open-ended process of interpretation and application, where societal values and norms can find (and usually find) a way of expression. The reference to fundamental rights, therefore, does not involve any closure to public involvement. They can rather be seen as "a public normative practice" (Beitz 2009: 170) in which the reference to fundamental rights do not exclude, but on the contrary implies the contribution of stakeholders and the public for determining their content and the concrete goals to be pursued.

Besides a general reference to safety as a paramount criterion for assessing technology and innovation, fundamental rights play a key role in assessing the ethical acceptability of the innovation process, representing "normative anchor points" characterizing the specific European approach to the ethical and regulatory challenges of innovation (Ozolina et al. 2012: 27), in particular with reference to the EU charter on fundamental rights (Ozolina et al. 2012: 27, van den Hoven et al. 2013: 58). Moreover, innovation is expected to take account of the societal needs "expressed in the Treaty on the European Union", as sustainable development, equality, quality of life (van den Hoven et al. 2013: 58). Yet, fundamental rights and societal needs are seen as explicitly and mutually linked goals of a comprehensive normative framework for the governance of science, technology and innovation.

In this sense fundamental rights are not simply constraints on innovation that aim to reduce or avoid its undesirable or negative consequences by warranting the respect for human health, dignity, privacy, etc. Rather, they also concern the shaping of policies, so that rights are not only respected and protected, but also promoted by way of proactive initiatives.

1.8 Concluding remarks

RRI can be deemed as a governance approach to research and innovation practices integrating fundamental rights and soft regulatory mechanisms and instruments. The efficacy of this approach is based on the combination of principle-based and outcome-oriented regulation. We emphasized fundamental rights as the main "building blocks" of principle-based regulation and, more in general, of this framework.

The combination of fundamental rights with soft and hybrid regulatory instruments seems particularly apt to cope with the situation to which RRI is called to answer. In the context of RRI, the reference to fundamental rights could be seen as an important component in the constellation of elements determining the ethical acceptability of innovation and techno-scientific developments.

The success of referring to fundamental rights as a solution to provide "normative anchor points" for RRI requires careful examination of the legal and regulatory framework in which STI activities are framed in the EU and, at the same time, a deliberate effort to construe a governance framework designed to ensure the complementarity between hard and soft regulation, legal norms and voluntary commitments.

This situation reflects the RRI focus on actors' *responsibilisation* and the appeal to their capacity of committing to some goals that are not mandated by law, under the perspective of a renewed approach to responsibility.

The potential of fundamental rights to successfully combine a stable normative orientation with openness and flexibility is ultimately a matter of how the basic requirements of the constitutional state can be preserved in the multilevel and manifold regulation that characterises RRI governance approach.