2 Locomotive breath? *Post festum* reflections on the EC Expert Group on Policy Indicators for Responsible Research and Innovation

Roger Strand and Jack Spaapen

2.1 Introduction: the alien phenomenon of RRI

The work of historians and archaeologists is very hard. With great effort they search archives, buildings, monuments and excavation sites for traces from the past. Worse, somehow they need to make sense of these traces of alien practices, cultures and institutions, inscribed with forgotten signs and languages.

Out of the many endeavours and events of the 21st century with which the historians of the future might struggle, this chapter will provide an early account of a particular episode in the development of the governance of scientific research and (in particular, technological) development. The episode on our mind is that which was called “Responsible Research and Innovation” – RRI – and more specifically, the attempts at making RRI a powerful concept in research and innovation (R&I) policy in the European Union (EU). One of these attempts was the creation of an expert group at the European level, called the Expert Group on Policy Indicators for Responsible Research and Innovation. The group worked intensively throughout 2014 and their final report was published in 2015 (Strand et al., 2015). This chapter is not at all a neutral and disinterested account of the work of this expert group; in fact, the authors were strongly involved in the work, as “rapporteur” (Jack Spaapen) and “chair” (Roger Strand), respectively.¹ Rather, we provide a first-hand narrative with some reflections on the results and the (lack of) translation of these results into policy and governance, at least to date, mid-2019. We have chosen to do so not only to enrich the archives for the historians of the future but also because we believe these reflections to be relevant for all those who work to develop the governance of R&I. From the perspective of 2019, it seems that although the concept of RRI might not survive for long – this we cannot know – the underlying structural challenges it was made to address will not go away any time soon. In other words, the future might be in the past when it comes to the basic ideas of RRI.

The blunt way of stating these structural challenges is to liken science to a runaway train, hijacked by policy and industry. At the beginning of the 21st

century, many countries and regions held economic policies that emphasized the role of science in relation to larger innovative processes. In many countries programmes were set up to connect scientific developments closer with industrial sectors in society and/or with wider societal challenges. In the Netherlands, the Top sector policy may serve as an example. Through this policy, nine economic sectors were identified that were deemed of vital importance for the Dutch economy, and a funding scheme was developed for intense collaboration between academic researchers and industry or public organisations (Government, s. a.). At the EU level, large parts of subsequent framework programmes were devoted to connecting science and industry and society in similar ways, culminating in the H2020 framework programme where, apart from the European Research Council section, all other parts of the budget are reserved for industry and the Grand Societal Challenges. Whereas scientific development had been taken as a hallmark of societal progress for at least two centuries (Rommetveit et al., 2013), it was mainly after the end of World War I (1918) that science became seen as an engine for economic growth (Godin, 2006). An important expression of this political imaginary was Vannevar Bush's (1945) report "Science: The Endless Frontier". Geographical frontiers, such as the Pacific Coast, had been conquered, thanks to explorers like Lewis and Clark, and other frontiers had to be given up at the entrance into post-war, post-colonial times. Bush's formula for conquering the "endless frontier" was that generous support to basic science would lead to technological inventions, new consumer products and therefore sustained employment and economic growth. At the turn of the 20th century, the economic policies of industrialised and wealthy nations emphasised not only being generous to science but also actively pushing research towards industrial innovation, increasing innovation rates and thereby keeping or gaining competitive advantage over other nations. In this political climate, Schumpeter's idea of "creative destruction" had been accepted and transformed into a maxim of accelerating and "disruptive" innovation, which was seen by governments as inherently desirable (notwithstanding the social side-effects of market disruptions).

The desire for the New was slowly contrasted with concerns about the New. Nuclear technology became controversial after the bombing of Hiroshima and Nagasaki, and later stayed controversial for different reasons after disasters with reactors such as in Three Miles Island (1979), Chernobyl (1986) and Fukushima (2011). More generally, concerns started to grow over the role of Science in other societal issues such as the defence industry, in industrial pollution, and from the 1980s, in engineering of life with the aid of molecular biology and biotechnology. Initial concerns were voiced outside of the political institutions and arenas and often labelled as anti-progress (religious, Luddite, irrational, pre-modern, New Age etc.). From around the 1980s, however, concerns also found expression in civil society which could be voiced inside the political institutions, notably as "ethics", "technology assessment", "risk assessment" and "public engagement" with science.

These different expressions co-existed and grew in force towards the end of the 20th century (Greenpeace and other non-governmental organisations concerned with the environment or the food industry or health issues) and became in many European countries part of the regular debate about the role of science in society, although in other countries remained more marginal in the policy discourse. They were somewhat supported, though, in a few national contexts and also through the minor “Science-and-Society” and “Science-in-Society” programmes within the EU framework programmes.

In 2011, the term “Responsible Research and Innovation” was introduced by philosopher and EU DG RTD (Directorate-General for Research and Innovation) policy officer René von Schomberg as he was summing up results from Science-in-Society research in FP7, the seventh EU Framework Programme for Research and Innovation. He defined RRI as:

a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products.

(von Schomberg, 2011, p. 9)

The term was soon adopted by DG RTD into its policy. By political intervention it was even made into a “cross-cutting principle” of the eighth framework programme, the so-called “Horizon 2020”. In von Schomberg’s initial mention, the potential of science and technology to cause harm was explicit:

we are confronted with the Collingridge dilemma, implying that ethical issues could be easily addressed early on during technology design and development whereas in this initial stage the development of the technology is difficult to predict. Once the social and ethical consequences become clearer, the development of technology is often far advanced and its trajectory is difficult to change.

(von Schomberg, 2011, p. 8)

Von Schomberg thus introduced the reference to Collingridge (1980) and his dilemma of the social control of technology into EU policy. Elsewhere, von Schomberg argued for the need for a new paradigm of governance. Faced with the accelerating runaway train of science and technology development, law, ethics and the market were all too slow and weak and they arrived late to the scene of the accident. Scholars such as von Schomberg argued that the new paradigm – whether called RRI or something else – could be created by joining the theoretical and practical forces of the various strands of science critiques – philosophy, ethics, feminist critique, studies of risks and hazards, participatory technology assessment and many others. The new paradigm was envisaged as “soft governance”: not strict laws – how can you regulate that which still is not discovered or even imagined? – but rather (more or

less) voluntary self-governance by scientists and technologists who would learn to deeply reflect on the prospects of doing good as well as harm to society and the world and who would also deeply engage in deliberation with citizens, not least those who might suffer from the potential harms. The paradigm was imagined in terms of more or less constructivist philosophies of science and technology. Science was not seen as a fully deterministic process that merely discovers the Truth about Nature, but rather as a human, social and cultural practice that could be shaped and influenced by politically motivated decisions. Helga Nowotny, Peter Scott and Michael Gibbons (2001) coined the term “socially robust knowledge”, referring to their earlier work on how the social contract of science had changed by the closer intertwining of science, industry and government. Socially robust knowledge, in their definition, would be knowledge that combined epistemic and social values – being scientifically sound, but also tested in societal practice. A similar concept of “post-normal science” was previously introduced by Funtowicz and Ravetz (1993).

When the European Commission (2012) adopted RRI as a policy principle, the result showed little evidence of regard to the Collingridge dilemma and the constructivist philosophies. Instead, it translated the intellectual traditions that formed the basis of RRI into operational tasks, or “keys” as they became known. Accordingly, the policy leaflets and explanations from the European Commission described RRI as something that could be verified by a box-ticking exercise:

1. Were there ethical issues, and were they taken care of in an appropriate manner?
2. Were there gender issues, and were they taken care of in an appropriate manner?
3. Did the R&I project engage with civil society?
4. Did it also educate the public to improve the quality of the social dialogue?
5. Did it provide open access of its results to the public?
6. Was there an appropriately responsible governance of the R&I project? (This “key” actually held the potential to redirect the attention to von Schomberg’s point and was mostly left undefined or even omitted.)

This was often called the “keys” approach to RRI – popular within DG RTD and by extension relevant to the many clients (including ourselves) who relied on the wealthy patronage of the Directorate-General to fund their research projects on and around RRI during the latter phase of FP7 and subsequently Horizon 2020. Still, within the intellectual RRI community that formed, von Schomberg’s definition and similar approaches, such as the “AREA framework” endorsed by the UK Engineering and Physical Science Research Council, were the ones considered in line with the original rationale of at least trying to deal with the runaway train. What developed

was a sort of subversive humanism, usually the mildest sort, in which the formal deliverables of FP7 and Horizon 2020 RRI projects complied with the five- or six-keys approach and delivered results on them, while the academic outputs – written and oral – took a freer stance towards the keys.

Sometimes the subversive humanism also included attempts at simultaneously embracing the key approach and the philosophical origin of RRI and somehow integrating them or expanding the former with the latter. The Expert Group on Policy Indicators for Responsible Research and Innovation was an example of such an attempt. Its mandate, as specified by DG RTD, was “to help the Commission identify existing indicators and to propose new indicators that can measure impacts of RRI activities in qualitative and quantitative terms” (Strand et al., 2015, p. 15). The terms of reference defined the aim of RRI as to “improve the societal alignment of R&I, in particular by improving public engagement, science education, gender equality, open access to scientific information, ethics and governance”, that is, in terms of the six keys. Furthermore, the mandate defined the need for indicators in terms of an underlying agenda of evidence-based policy-making, demanding that indicators be “SMART” (an acronym which in this particular context meant “specific, measurable, attainable, relevant and timely”). In other words, the mandate was quite far from von Schomberg’s vision of philosophically informed self-governance among researchers who sought reflection and deliberation in civil society. Our group went to the task with the same mild brand of subversive humanism as could be observed in many of the RRI/Science-with-and-for-Society projects conducted in the same period. We did propose indicators, and they did address the six keys. However, our indicators were hardly SMART in the above sense and in few cases quantitative or easily measurable for a policy directorate. Indeed, we politely opposed the agenda of evidence-based, top-down command-and-control style of governance and rather proposed a large suite of indicators together with an approach to selecting among them that re-introduced the ideal of self-governance. In that sense, our approach was “smart” in the street-wise sense of the word; we acknowledged the contextual character of R&I by presenting a toolbox of indicators, mostly qualitative, from which the community of researchers and stakeholders could choose those relevant for their project. The context of health research is completely different from the context of global warming or migration or electromagnetic radiation by mobile telephone masts. And it is the community of researchers and stakeholders that is best to choose the most relevant indicators. Our Expert Group therefore chose a slightly more subversive strategy than the concurrent Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) project, which actually produced measurable indicators that apparently could work in a command-and-control type of governance system. In doing so, we could be seen to find inspiration in the report of a previous expert report within the same area, called “Taking European Knowledge Society Seriously” (Felt et al., 2007). That report

could hardly be seen to answer the operational questions it was asked and hence got little direct policy impact in the short term. Its depth of analysis, however, had made it somewhat of a classic within the academic community and still a source of inspiration a decade after it was produced.

2.2 The results from the Expert Group: a suite of indicators for a network approach

Even outside the logic of control-and-command, the question of monitoring and devising indicators can make good sense. Also in bottom-up self-governance it is reasonable to ask if a certain action or activity was successful and met the purpose it was intended to fulfil. To reflect upon one's criteria for success is worthwhile irrespective of model of governance. In the case of RRI, there have been various proposals for RRI indicators. The outcomes of the MoRRI project are described elsewhere in a wealth of documentation.² Wickson and Carew (2014) were not obliged to stay with the six keys and presented an indicator approach closer to the academic literature on RRI.

As mentioned above, the Expert Group chose to try to integrate the “philosophical” or “intellectual” concept of RRI, based on von Schomberg's definition, with that of the six keys. The former was given priority in the sense that it gave direction to the choice of monitoring framework and the content of indicators. For instance, the Expert Group interpreted “ethics” as a key of RRI in terms of those issues and aspects of RRI (philosophically defined) that pertained to ethics or happened to be actualised in practices and institutions with the ethics label. Following this argument, the ethical acceptability of R&I outcomes was seen as the main ethics dimension of RRI, whereas the typical issues of medical research ethics such as informed consent fell outside the scope. How such issues (such as informed consent or more generally ethical clearance) were dealt with institutionally, however, and by which mechanisms, could be an RRI ethics issue. Similar arguments were carried out for the other keys.

Giving the philosophical concept of RRI priority also implied the need to deal creatively with the idea of a monitoring framework to make it compatible with the appropriate understanding of governance: not command-and-control but an appreciation of the role of new and broader dialogues between science and civil society. The approach chosen was largely inspired by the concept of productive interactions developed in the FP7 SIAMPI project, referring to societal change being the consequence of interactions between researchers and stakeholders in society, defined as a dynamic network around a societal issue like, for example, global warming, renewable energy or migration (Spaape and van Drooge, 2011).³ This was stated in the executive summary of the Expert Group report as follows:

This means that the group sees RRI from a network perspective, consisting of stakeholders jointly working on a set of principles guided by the RRI keys. When monitoring RRI policy, the prime focus therefore

should be on the development of RRI agendas in these networks. In other words, the focus of monitoring and indicators should be on the governance of RRI in these networks, and in particular on what kind of RRI policies are developed (best practices). Secondly, one should discuss the kind of indicators best fitting research and innovation (R & I) practices and goals within these networks, and then decide upon a limited set of indicators (because things have to be manageable). Such indicators are thus highly contextual, meaning that there will be no one list for all.

(Strand et al., 2015, p. 5)

The strategy of the Expert Group was accordingly to propose a large set of more or less practicable indicators that could give ideas and inspiration not only to the European Commission but to anyone who might be interested in monitoring the success of RRI activities in a learning process.

2.2.1 What to indicate?

Another design issue, perhaps a less controversial one, was the question of what the indicators were supposed to indicate. RRI activities and initiatives could be seen as a means to an end, namely to change R&I practices so that they would be more “responsible” in the RRI sense. In principle, one might have a theory or a body of evidence that already specified what such means would have to be in order to reach that end. As a matter of fact, one did not. The causality between RRI means and RRI outcomes was not established. On the other hand, measuring the degree of responsibility of R&I practices was not a trivial exercise, either. The Expert Group opted for a hybrid model:

To be responsible in general and in the specific terms of RRI includes three dimensions: performance, perception and key actors. RRI performance depends both on the processes that promote RRI activities and on the effects that these processes have: outcome. Acting responsibly defines who we are – we are acting in a certain manner (performance). But responsibility also includes the key element of perception – to be seen to act responsibly. The weight given to process, outcome and perception indicators for a given issue depends upon the nature of that issue. Again, what we offer here is a generic and flexible framework that in itself can be applied responsibly as well as irresponsibly.

(Strand et al., 2015, p. 16)

2.2.2 What about substantive values such as sustainability and social justice?

In addition, there was the issue of the comprehensiveness (or rather the lack thereof) of the six keys. In line with the strategy of “mild subversive humanism”, the Expert Group kept the keys but presented an argument for extending the

framework with two additional keys: sustainability and social justice/inclusion. Although the argument never gained any traction, at least not until the writing of this chapter, we took some pride in how it was crafted. It was presented as a corollary of the Europe 2020 strategy of the EU, an overarching policy of which actually Horizon 2020 was an act of implementation. The slogan of Europe 2020 was “smart, inclusive and sustainable growth” (Strand et al., 2015, p. 11). The Expert Group argued that, if R&I was to give growth, then inclusive and sustainable R&I was needed to yield inclusive and sustainable growth. While perhaps not an entirely valid syllogism, the argument was expressed within the orthodox EU policy discourse at the time; within what a Rawlsian would have called the comprehensive normative doctrine.

The resulting framework could be presented as a matrix (Table 2.1).

The remaining task was to populate the cells. The Expert Group did so with creativity, with little regard for the statistics that were already sitting there in DG RTD, in part because we expected the MoRRI project to take a more pragmatic approach, and in part because the policy winds inside DG RTD were blowing away from, or even against, RRI towards the end of 2014. As it became ever less likely that the indicators would actually be adopted by DG RTD in their daily work, it appeared ever more important to present something that could inspire RRI governance “out there”, in national research funding organisations and in self-governance in research-performing organisations. Again, the ethics indicators may serve as an example of the creative approach. The following were the proposed indicators for the process dimension of the RRI ethics key (excerpts from report) (Strand et al., 2015, pp. 34–35):

- *Documentation of institutional attention to normative tensions related to research integrity policies and actions. (This may manifest itself in a*

Table 2.1 The indicator framework of the Expert Group report

Criteria	Performance indicators		Perception indicators	Key actors
	Process indicators	Outcome indicators		
Public engagement				
Gender equality				
Science education				
Open access				
Ethics				
Governance				
Sustainability				
Social justice/inclusion				

Source: Reproduced from Strand et al. (2015).

number of ways, e.g. activities such as open meetings or seminar series on ethics and research integrity and the inclusion of such discussions in research strategy papers, research proposals.) This is a qualitative indicator that will have to be based on strategic sampling of information – process indicator.

- *Documented change in R&I priorities (research or research funding) attributable to multi-stakeholder and/or transdisciplinary processes of appraisal of societal relevance and ethical acceptability (presence/frequency; qualitative descriptions; best practices) – process indicator.*
- *Presence of mechanisms for multi-stakeholder and/or transdisciplinary processes of appraisal of societal relevance and ethical acceptability (presence/frequency; qualitative descriptions; best practices) – process indicator.*
- *The formal and actual scope of the ethics review/IRB [Internal Review Board] clearance (e.g. whether a committee or a reviewer is free to identify any type of ethical issue or concern or has to limit the analysis to a predefined set, and the degree to which this freedom is enacted) – process indicator.*

The first entry of this list illustrates again how the “philosophical” concept of RRI informed the design. Research integrity as such might be an issue of research ethics but, in this understanding, not of RRI. Rather, the RRI issue is whether the R&I actors and practices are aware of the tensions between the theory and practice of research integrity, which again was seen to be related to the tensions between Mertonian norms and a conception of science as a fully autonomous, ivory tower-type institution on one hand, and the social contract of science that was envisaged by the more constructivist or at least coproductionist views on science that informed RRI thought, on the other. To be concrete, installing and applying ethics guidelines for co-authorship is a research integrity issue. To reflect and act upon the tension between such ethics guidelines and the real-world pressures and demands for strategic authorship in a highly competitive research field is an RRI issue. Even just to understand what this indicator referred to, not to speak of how to measure it, demanded insight into these theoretical issues. But why should such insight not be demanded? Why would it be appropriate to allow people to govern science if they have no theoretical understanding of the process of how science is organised and carried out?

2.3 What does it mean to take a “network approach”?

The preceding paragraph ended with two rhetorical, tendentious questions. Of course we stated them because we believe that it is *inappropriate* that science is governed by individuals and institutions who lack a theoretical understanding of science. Such an understanding cannot be gained in a short while or by a short text; a proper learning process would require the study

of STS scholars such as Sheila Jasanoff, Bruno Latour, Donna Haraway and Brian Wynne (to mention a few) in addition to the already-mentioned Nowotny, Gibbons, Funtowicz, Ravetz, Godin and others. It is only fair, however, that the authors of this text try to indicate what we mean by a “network approach” and a “coproductionist perspective” to those readers to whom such concepts are alien. We shall do that by providing our own, possibly somewhat idiosyncratic, outlook on the governance of science at the time of writing.

At the beginning of the 21st century, in countries with a functional research sector, institutes operating within the academic sphere (universities, academy institutes and other) also perform their research in the context of issues and challenges in policy or society at large. The networks that are formed by these wider communities are as a rule more diverse and often also more temporary than traditional academic networks. Both researchers and policy makers have to review their perspective on this position, both in terms of how they interact with the broader environment (for example, with regard to their R&I agenda) and in terms of how they assess the success of new forms of collaboration. The networks are characterised by a variety of scientific stakeholders (various disciplines) and stakeholders from society, be it industry, government or society at large. Somehow, all these different backgrounds, interests and work practices have to be attuned in new arrangements in which goals, performance and monitoring and evaluation have to be elaborated.

Research that aims at helping innovate society in a responsible way, whether this is through new ideas, technologies or new forms of organisation, new insights or processes, or otherwise, is thus mostly the outcome of an intricate process in which a multitude of actors operate, with participants coming into the collaboration with different perspectives, interests and expectations. The mix of partners differs per sector and per subject, and so will the weight of different RRI keys. Sometimes there is a small enduring network of established actors, with a high level of coordination, for example in public health research where researchers collaborate with specific governmental bodies, hospitals, general practitioners and patient organisations, and sometimes industry (Spaapen and van Drooge, 2011, p. 214). Typically, the research agendas in such networks are developed in close collaboration with the relevant stakeholders who have more or less the same goal. It is relatively easy to agree about which RRI factors are relevant and how they should be applied. In other cases, this is not so. Networks can be much more open and dynamic, with stakeholders coming in and going out, and having rather different goals in mind. One example is the case of energy transition, where goals and development paths can lie far apart (nuclear energy versus renewable energy) and stakeholders feel less connected to each other (say, the car industry and Greenpeace). But also in those cases, RRI should be part of the agenda development. As we write, in 2018, *automated cars* are emerging as an issue of growing importance and potential for controversy. We

shall take that issue as an example of a diverse network filled with different interests.⁴

The automobile industry in Europe and the USA is starting to invest in the development of electrical and self-driving cars, from Ford's \$1 billion investment in an artificial-intelligence company to Uber's 2016 purchase of self-driving truck company Otto. Tesla, Nissan and BMW all say they will have fully driverless cars by 2021. And a report from the Boston Consulting Group suggests that by 2030 more than 5 million conventional vehicles could be replaced by automated ones in the USA alone (Lang et al., 2017).

Many European governments have set up policies to ban fossil fuel cars in the next two decades or less. Obviously, a main driver for this development is the necessity to reduce the level of CO₂ reduction in line with the Paris agreements. But the new transportation system will have a major influence on many other aspects of society. It will be a relief to drivers who have to commute every day from home to work or elsewhere, and then it will most likely influence the behavioural patterns of these drivers. To predict what will happen is however very difficult since, as it turns out, the environmental impact of automated vehicles is extremely uncertain. One report from the American Department of Energy found that automated vehicles could reduce fuel consumption for passenger cars by as much as 90%, or increase it by more than 200% (Worland, 2017).

The wide range of potential outcomes is the result of a long list of variables about how a future with automated cars will take shape. These variables partly lie in the scientific and technological options that will be developed, and partly in the behaviour and interests of a multitude of societal stakeholders, ranging from industry to governments to consumer organisations to the general public. Many researchers expect that automated cars will lead to a sharp increase in the average miles travelled by a given vehicle, thanks to the human factor. Key barriers to hopping in a car – fatigue, age or intoxication, to name a few – will disappear, they expect, and car owners will be free to travel further and more frequently. Workers may choose to live even further away from the office, opting to sleep in the car or use that time to squeeze in a workout. And, once in the city, car owners might instruct their vehicle to drive around in circles rather than pay for parking. And there is more. Beyond changed driving patterns, simple technology advances will reduce the environmental toll of automated cars. Most importantly, engineers say that the largely accident-free vehicles can eliminate safety equipment, such as antilock brakes and airbags, that has increased the weight – and fuel consumption – of vehicles. Automated cars can also travel closer together, allowing them to take advantage of aerodynamics. Trucking fleets are already trying to take advantage of this fuel-saving measure.

With all the insecurities the question is how to (try to) steer – *govern* – this important development forward. Who are the key players and how will they relate to each other, and will they develop, and carry out, an R&I

agenda in a responsible way? Researchers alone cannot do this. They have sought to model how humans might respond to automated driving using surveys, driving data and lab experiments, but ultimately the sheer number of choices and assumptions involved in transportation has made reaching concrete conclusions about driving behaviour difficult. Regulators cannot do it. While they represent the obvious way to protect against the potential environmental downsides of automated vehicles, they will not know what the most efficient way is to influence drivers' behaviour as long as they are uncertain about technical and social developments. The automobile industry cannot make the transition without changing their business model that historically has been founded on fossil fuel cars. Clearly, the network around this issue is much wider (for example, how will enough electricity be produced without negative effects on the environment, and by whom?), and the actors in this network will be confronted with RRI issues on a more or less permanent basis.

To sum up: there are many uncertainties in the scientific and societal developments of automated cars and road systems. The future will be decided via a mix of technical solutions, human behaviour and political and business decisions. Many RRI keys play an important role in many of the discussions between the different stakeholders: how to engage civil society, educating the public to strengthen social dialogue, ethical and gender issues. But perhaps most of all the question how to govern such a process in a way that is inclusive and socially just. While scientists and policy makers say reducing emission figures will be key to addressing man-made climate change, there is a wide range of potential outcomes and this is the result of a long list of technical, political and human variables about how a future with automated cars will take shape.

We see an area where there seems to be a lot of consensus about the necessity of reducing the emission of CO₂ as a prime cause of global warming and climate change. But at the same time, there is a lot of uncertainty how to reach those goals in a responsible way in a societal network with lots of different interests and uncertainties. The main challenge is to develop a process of development that is based on only responsible governance. Arguably, this is more difficult to manage in the latter case than in the case of public health, mentioned before, where we see a more or less stable network with partners close to each other.

2.4 RRI: a sinking ship?

Whether such a responsible development will be possible in the USA of the early 21st century remains to be seen. For those who still believe in so-called European values, as professed in the EU Charter on Fundamental Rights, including the present authors, it seems that RRI ought to be part and parcel of European development. Listening to policy makers in Brussels, or reading the research policies at the entry into the ninth Framework Programme for Research

and Innovation, however, RRI seems to be a sinking ship, if not already at the bottom of the Atlantic ocean. Still, if Europe wants to be different from the rest of the world, more democratic, more social, more ethical, etc., then it is difficult to see how one could avoid the values and principles informing RRI in one way or another as part of the development of R&I in a responsible society. To do this effectively, a new thinking about agenda setting and evaluation of network activities is necessary. In short, this should be based on the combination of a theory of change and a convincing narrative to explain both the main goal of R&I and the ways how to achieve these goals.

Evaluation of R&I in the 21st century is therefore both demanding and exciting. It is exciting because the pace of scientific research speeds up by the year thanks to the growth of science and global collaboration. Adding to that is the influence of digitisation, enabling new ways of cooperation, and collecting data and exchanging them globally, the growing need for inter- and transdisciplinary collaboration (especially between human and technological scientific fields as machines and human bodies are becoming ever more connected; see Strand and Kaiser, 2015)) and the worldwide challenges that are in desperate need of scientific input (climate change, energy transition, genetic engineering, migration and political crises). One could say that science is finally becoming a truly international trade.

It is demanding too, because a broad variety of stakeholders (from policy, industry and wider society) is pervading university life at an ever faster pace. Academic researchers are required to think about the relevance of their work for those stakeholders, and need to engage in a negotiation process with them. There is no escape here, because more and more research funding is allocated in ways that require societal impact, both at the national and European level. Europe spent about 80 billion euros in the current framework programme H2020 and two of the three main parts – the Grand Societal Challenges and industrial leadership) directly promote collaboration between science and society. At the national level governments design large programmes to stimulate collaboration between science and society, like, for example, in the Netherlands the Top Sector policy and the Science Agenda (Dutch National Research Agenda, s. a.). In this, demands for societal impact and public engagement come together, the Dutch National Research Agenda being the result of a nationwide public consultation about what would be the most urgent questions that science could help solve in society. Solutions will have to come from collaboration between researchers and variegated stakeholders in society. To assess these kinds of large multi-partner network activities, the Dutch have a comprehensive evaluation method available, the Standard Evaluation Protocol (SEP).

2.5 ... or a Flying Dutchman?

The national evaluation framework for publicly funded research in the Netherlands may serve as an example of comprehensive evaluation which

Table 2.2 Overall framework

	<i>Scientific quality</i>	<i>Relevance to society</i>
Demonstrable output		
Demonstrable use		
Demonstrable recognition		

includes RRI-like elements. Operating since 2003, the SEP is now in its third version, running from 2015 to 2021. The philosophy behind the SEP is that: (1) evaluators should balance the assessment of scientific quality and societal relevance; and (2) SEP offers an overall framework for all scientific fields, but within that all fields should be able to use indicators that are appropriate for the research practices of a particular field. In other words, there is room for a bottom-up development of indicators as long as they fit into the overall framework. The overall framework consists of two main criteria, scientific quality and societal relevance, with three indicator categories for both: output, use and recognition (Table 2.2).

Next to these two main evaluation criteria, SEP includes a number of other criteria that are part of the overall assessment framework. These are viability, PhD programmes, diversity and research integrity. With the exception of PhD programmes, all of these criteria overlap with the RRI keys set by the EU. Viability is meant as a management criterion which includes governance issues, diversity includes gender issues and research integrity includes ethical issues.

Research entities to be evaluated have to produce a self-evaluation report that regards all of these elements, preferably via a narrative highlighting the connection between the scientific goals and their relevance to society. Indicators may be used to underpin this narrative, but, as mentioned above, the SEP does not prescribe which indicators to use. There are two important ideas behind this: one is that there is not one set of indicators which is useful for all fields, and the other is that the research community should be allowed to take responsibility by choosing how to represent its research production and the ways it interacts with stakeholders in society. The main difference with other evaluation systems is that, in the SEP, the responsibility is placed with the community of researchers and stakeholders to enable self-governance.

RRI, in von Schomberg's (2011) definition, fits in this approach of comprehensive evaluation. As mentioned above, some of its main keys have been introduced in the national evaluation process, SEP 2015–2021, like diversity/gender issues and ethics/integrity.

To assess research in this broader context, rather different evaluation regimes than the current ones are necessary. Innovative approaches can be found in a limited number of countries, for example the Scandinavian

countries, the UK, Canada, Australia and the Netherlands. Building on the possibilities that SEP 2015–2021 offers, the humanities faculties have developed a novel approach which is based on the *narrative* as a central element in the self-evaluation report. The idea is simple: tell the world why it is important that you are doing the research you are doing. To do this in a convincing way for a wider audience, you need to clarify the connection between the scientific goals and the possible impact results can have on society. And you need to show how you envisage engaging stakeholders and how you collaborate with them. To be short, research is part of a bigger innovation process, and its role in that process has to be shown.

In principle, this is a qualitative approach, but the narrative should be supported by concrete evidence of activities in the realms of science and society. The purpose of this new kind of evaluation is different from traditional evaluations: it is not competitive and ranking-oriented (who publishes the most), but oriented towards collaboration and mutual learning: how can we best bring together necessary knowledge and expertise from science and society to solve major problems in the world?⁵

2.6 Off track and into the sea

We can imagine how the historians and archaeologists of the future might conclude how governance of science emerged as a particularly difficult societal challenge in the 20th and 21st centuries, characterised by high but unclear stakes, partial understanding, disputed criteria and sometimes outright confusion. Metaphors have their benefits and risks when understanding is partial. Metaphors can help to give a first, tentative grasp of the situation but may also direct the attention too much and effectively limit it. “Governance” is already, by its etymology a nautical metaphor, indicating that science is a ship to be navigated. It is not, if at all, on the sea, it is a huge fleet, and in a sense it is not too wrong to think of the marginal attempts of imposing RRI as top-down policy concept as the attempt to launch another little ship into that fleet. Playing on that metaphor, it was little wonder that it sank, and better prospects of success if it could become a network approach Flying Dutchman, that is, a ghost that might haunt, fascinate and inspire the minds of the thousands of sailors and fellow travellers who together constitute the collective governance of science of the future. Whether that ghost will carry the name RRI is not necessarily as relevant. The underlying issues of difficult collective choices will remain.

Indeed, in order for human historians and archaeologists to exist in the future, the governance of science will have had to improve. Initially in this chapter, we likened science to a runaway train, hijacked by policy and industry. This is an appropriate metaphor in order to emphasise the seriousness of the situation. Indeed, just by compiling a list of some of the products of science and engineering in the 20th century – nuclear technologies, plutonium, chemical and biological weapons, surveillance technologies, genetic

engineering of bacteria and virus, increasing CO₂ levels, microplastics, organic pollutants, algorithms that know nothing but pretend to know everything about us, etc. – it seems quite surprising that human civilisation did not collapse already in that century (or is the downfall happening now with some delay?). How much more can it take of irresponsible technological development?

The problem with the metaphor of the runaway train, however, is that it does little but consolidate a sense of disempowerment, as in the Jethro Tull rock song that gave the title to this book chapter, *Locomotive Breath*.

To create support for expansive research funding, 20th- and 21st-century policy makers crafted their policies within the discourse of technological optimism and scientific determinism: science is a train, its destination is Utopia and we should simply decide to go faster. As the Utopia becomes dystopic, optimism flips into pessimism and despair, as in the song lyric.

However, science is not a train, there are no rails and there is no predetermined trajectory. To admit that is the first step towards a more proper governance. The next step is to admit that there is no captain but rather a fuzzy, diverse, self-organised network of sailors and travellers with sometimes rather different visions of what can or should be done (think of the Chinese researcher He Jiankui and his genetically designed twin babies: Regalado, 2018). We hypothesise that the future historians who encounter this text will find these claims to be self-evident, otherwise they won't be there to encounter it. For them, then, the alien character of the early 21st century was that it was a time when these simple insights were controversial and could only to a very little extent be adopted into governmental policies and governance practices.

Notes

- 1 Our observations and reflections are entirely our own and should not in any way be taken to represent the views of other members of the mentioned expert group or individual fonctionnaires of the European Commission, nor seen to express the official views of the European Commission at any point in time. One author (Strand) has performed the work with this text as part of the project SuperMoRRI. This project has received funding from the European Union's Horizon2020 Research and Innovation Programme under grant agreement no. 824671. We are obliged to clarify that the Funding Agency cannot be held responsible for any use that may be made of the information the text contains.
- 2 All MoRRI deliverables can be accessed from <http://www.morri-project.eu>. The full indicator report, called Deliverable 4.3, can be found at http://www.technopolis-group.com/wp-content/uploads/2018/02/D4.3_Revised_20022018_clean.pdf
- 3 The FP7-funded SIAMPI project developed indicators to assess the social impact of research. More information can be found on the project website: www.siampi.eu
- 4 This example is adapted from an article by Justin Worland in *Time* of 16 November 2017, Why self-driving cars might not lead to a huge drop in fuel consumption.

- 5 The website for this approach was developed for the humanities, but may serve as inspiration for all fields: www.qrih.nl/en

References

- Bush, V. (1945). *Science: The Endless Frontier*. United States Office of Scientific Research and Development, U.S. Government print office, Washington, DC.
- Collingridge, D. (1980). *The Social Control of Technology*. Pinter, London.
- Dutch National Research Agenda (s. a.) *Dutch National Research Agenda*. Retrieved from: <https://wetenschapsagenda.nl/?lang=en>.
- European Commission (2012). *Responsible Research and Innovation: Europe's ability to respond to societal challenges*. DG Research and Innovation. DOI:10.2777/95935
- Felt, U., Wynne, B., Callon, M., Gonçalves, M., Jasanoff, S., Jepsen, M., Joly, P., Konopasek, Z., May, S., Neubauer, C., Rip, A., Siune, K., Stirling, A. and Tallacchini, M. (2007) Taking European knowledge society seriously. Report of the Expert Group on Science and Governance to the Science, Economy and Society Directorate, Directorate-General for Research, European Commission (EUR 22700). Retrieved from https://ec.europa.eu/research/science-society/document_library/pdf_06/european-knowledge-society_en.pdf.
- Funtowicz, S. and Ravetz, J. (1993). Science for the post-normal age. *Futures*, 25: 739–755.
- Godin, B. (2006). The linear model of innovation. The historical construction of an analytical framework. *Science, Technology, & Human Values*, 31 (6): 639–667. DOI: 10.1177/0162243906291865.
- Government (s. a.) *Encouraging innovation*. Retrieved from www.government.nl/topics/enterprise-and-innovation/encouraging-innovation.
- Lang, N., Rüßmann, Chua, J. and Doubara, X. (2017) *Making Autonomous Vehicles a Reality. Lessons from Boston and Beyond*. The Boston Consulting Group, Boston, MA.
- Nowotny, H., Scott, P. and Gibbons, M. (2001). *Re-thinking Science*. Polity, Cambridge.
- Regalado, A. (2018, November 25). Exclusive: Chinese scientists are creating CRISPR babies. *MIT Technology Review*. Retrieved from www.technologyreview.com
- Rommetveit, K., Strand, R., Fjelland, R. and Funtowicz, S. (2013). *What can history teach us about the prospects of a European Research Area?* (EUR 26120). Retrieved from <https://ec.europa.eu/jrc/en>
- Spaapen, J. and van Drooge, L. (2011). Introducing ‘productive interactions’ in social impact assessment. *Research Evaluation*, 20 (3): 211–218. DOI: 10.3152/095820211X12941371876742.
- Strand, R. and Kaiser, M. (2015). *Ethical Challenges Report on Ethical Issues Raised by Emerging Sciences and Technologies*. Council of Europe, Committee on Bioethics, Strasbourg.
- Strand, R., Spaapen, J., Bauer, M. W., Hogan, E., Revuelta, G., Stagl, S., Paula, L. and Pereira, Â. G. (2015). Indicators for Promoting and Monitoring Responsible Research and Innovation. Report from the Expert Group on Policy Indicators for Responsible Research and Innovation. (EUR 26866). Retrieved from http://ec.europa.eu/research/swafs/pdf/pub_rri/rri_indicators_final_version.pdf

- von Schomberg, R. (2011). Introduction. In R. von Schomberg (Ed.), *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields* (pp. 7–15). Publications Office of the European Union, Luxembourg.
- Wickson, F. and Carew, A. L. (2014). Quality criteria and indicators for responsible research and innovation: learning from transdisciplinarity. *Journal of Responsible Innovation*, 1 (3): 254–273. DOI: 10.1080/23299460.2014.963004.
- Worland, J. (2017, November 16). Why self-driving cars might not lead to a huge drop in fuel consumption. *Time*. Retrieved from www.Time.com