# **Reckoning with assessment: Can we responsibly innovate?**

Emad Yaghmaei and Ibo van de Poel (eds.): Assessment of responsible innovation: Methods and practices, 1st edition. London: Routledge, 2020, 394 pp, £96 HB

Steven Umbrello

Steven.umbrello@unito.it

Institute for Ethics and Emerging Technologies

Università degli Studi di Torino

Torino, Italia.

*Assessment of Responsible Innovation* argues, contrary to common imagination, that the profit motive underpinning private sector decision-making about innovation neither excludes—nor is even necessarily in tension with—responsible innovation. Responsible innovation is not a clear-cut thing, principle, or clearly formulated grouping of practices. Rather, it consists in a plurality of engagements, strategies, and interactions oriented around the general goal of technological development towards socially desirable ends. The assessment of responsible innovation faces a lacuna partly due to this plurality, and partly because responsible research and innovation (RRI) has primarily been the domain of research institutions, higher education, and public sector entities—those who are not responsible for the majority of innovations. There is thus a gap between past RRI research and the actual nexus of innovation programmes.

The editors are Emad Yaghmaei, Senior Researcher at the Faculty of Technology, Policy and Management at Delft University of Technology, and Ibo van de Poel, Anthoni van Leeuwenhoek Professor in Ethics and Technology and Head of the Department of Values, Technology and Innovation at TU Delft. They bridge the interdisciplinary contributions of this volume to offer a more cohesive and up-to-date understanding of how to assess RRI principles and practices. Their goal straddles multiple levels of abstraction: they not only evaluate the placement of contributions in the volume as a function of filling the gap in the assessment of RRI projects, but also assess RRI principles and align overall assessment practices with the United Nations (UN) Sustainable Development Goals (SDGs). This speaks to the larger editorial project, which is to ensure that innovations do not simply avoid doing harm but actually contribute to the social good (c.f., van de Poel 2020). As such, the volume seems to fit neatly into the wider corpus on RRI and the ethics of technology, which has distinguished TU Delft and its partner institutions as an epicenter of RRI research.

Yaghmaei and van de Poel divide the volume into three parts. The first two are primarily contextual, with the first shortest part discussing RRI more generally to set the stage for what proceeds. The second section frames RRI within the context of industry, arguably the nexus within which most innovation takes place, and an area that has only more recently become the focus of RRI projects. Given that RRI has traditionally received attention from university and public sector organisations, it has remained functionally extricated from the domain that the majority of such innovation could benefit. The core of the volume emerges in the third section, which is comprised of chapters laying out methods and approaches for the assessment of responsible innovation practices.

In their introduction, the editors recognise the need for a more grounded approach to understanding how RRI is conducted in the real world, hence the inclusion of chapters on RRI in the private sector. Two of the four chapters in this section focus on the food industry in particular, which speaks to the specificity and situatedness that RRI projects can take. But this is not in lieu of more generalised, aggregate accounts of assessing RRI in industry. In the fifth chapter, Porcari et al. use a multi-criteria approach to analyse eight European Union-funded RRI projects in industries with the aim of determining the costs and benefits of such an explicit RRI approach. The authors speak to the larger issues facing RRI in the private sphere, namely, that the benefits of RRI are not always *prima facie* obvious whereas the costs of implementation are often more imminently manifest. They contend that if one takes the implementation of RRI practices within a single company as the object of analysis, rather than the traditionally more abstract commentary on the benefits of RRI, then the individual benefits of such an approach for any given company become clearer. The authors argue that through this means, a more general strategy can be developed and implemented to reveal the boons of RRI implementation.

The final section is composed of seven ‘best practices’ for RRI assessment, and this is where the volume makes its mark. Given that the most common shortcoming of edited collections is incongruencies between contributions, the volume’s deliberate buildup from context to practice could have been especially anticlimactic. But here issues of chapter cohesion does not appear problematic as each chapter is buttressed by the underlying tensions of RRI assessment more generally. For example, Andrea Porcari and Elena Mocchio’s best practice #5 formalises how to manage the social impacts and ethical issues of RRI, providing both the philosophical and methodological substrata for the subsequent best practice. Best practice #6, by Lisanne Urlings, outlines how to benchmark corporate performance towards higher-order goals of innovating *for* the UN SDGs. Much of the cohesion here points to, not only the intimacy of the authors’ understandings of RRI literature at large, but their mutual familiarity and coordination. Despite their formulation of tools and practices for assessing and benchmarking RRI boons and shortfalls, the authors of both best practices affirm the necessity of modularity for these tools. This speaks to the more general philosophical substrata of RRI: the contextual and social situatedness of technological design as fundamental to how innovation is carried out. Overly general tools that are incapable of adaptation to particular settings will be impotent.

Van de Poel’s concluding chapter speaks to this issue more succinctly, distinguishing issues in RRI assessment as falling into categories of (1) the measurement of RRI, and (2) settings for RRI assessment. Van de Poel argues that these categories of pitfalls are not mutually exclusive. They can feed into each other, reinforcing the temptation of industry to engage with RRI in only a cursory way rather than fundamentally adopting the principal motives of RRI. To address the issue of measurement, van de Poel maintains that the underlying rationales for engaging in RRI practices need to more explicitly teased out. be expounded. Such rationales are often in tension with one another and, as the volume as a whole suggests, specification is key. The road to successful RRI assessment involves choosing one rationale over trying to accommodate all of them.

RRI assessment remains a fascinating feature of what is now more canonical work on the implementation of RRI practices in specific domains and projects. The volume is aimed at an interdisciplinary audience, mirroring the interdisciplinarity of its contributors. In fact, most chapters explicitly cite work outside, albeit related to, the body of RRI literature, drawing from engineering, design, and the philosophy of technology. Owing to this, the volume should attract multiple demographics. But given the jargon-laden style of most of the chapters, its appeal will be constrained to specialists in those domains. If time is taken to attend to these dense contributions, then seasoned scholars will find their labour rewarded. Ultimately, Yaghmaei and Van de Poel have succeeded where most stumble: they have managed to create a congruent edited collection focused on a burgeoning field of study that reads more like a cohesive monograph than a collection of eclectic contributions. While it may only be attractive to a well-read audience on the topic, it is a strong primer for graduate student teaching on approaches to assessing and measuring RRI project success. The effort comes highly recommended.

**References**

Van de Poel, I. 2020. Embedding Values in Artificial Intelligence (AI) Systems. *Minds & Machines* 30, 385–409. https://doi.org/10.1007/s11023-020-09537-4