4TU.CEE

BEFORE RESPONSIBLE INNOVATION: TEACHING ANTICIPATION AS A COMPETENCY FOR ENGINEERS

T.W. Stone[1]
Delft University of Technology
Delft, The Netherlands

L. Marin
Delft University of Technology
Delft, The Netherlands

J.B. van Grunsven
Delft University of Technology
Delft, The Netherlands

Conference Key Areas: *(3) sustainability and ethics; (9) future engineering skills and talent management*
Keywords: *engineering ethics, Responsible Innovation, anticipation, virtue ethics*

ABSTRACT

This paper focuses on engineering ethics education utilizing Responsible Innovation (RI). As a forward-looking approach aiming to embed ethics within innovation practices, RI strives to align technology development with societal values. However, when teaching the concepts and methods of RI, we face two intertwined challenges. First, RI presupposes we can estimate the consequences of an innovation or design intervention, while evidence shows it is nearly impossible to fully predict the consequences of new technologies. RI acknowledges this by replacing an ambition to *predict* with a call to *anticipate* innovation-consequences. However, without a robust account of anticipation this merely kicks the can down the road. Second, RI seems to suggest that we know what is meant by a specific value (e.g., privacy, sustainability) and its relation to a specific technology. While such knowledge is key to an anticipatory perspective, values are often treated superficially and a-historically in RI literature. To address these challenges, we argue that RI-focused education – and engineering ethics generally – should be fostering *historically informed anticipation* as a core competency. To do so, we will define and characterize a set of interrelated virtues essential for engaging in historically informed anticipation: *moral sensitivity* (an ability to identify values at stake), *epistemic humility* (an awareness of the limits of one's understanding), and *moral imagination* (an ability to envision new perspectives and solutions). We suggest this can be cultivated via a novel teaching method that involves an in-depth historically informed normative analysis of a value-technology dynamic (called a *value-genealogy of technology*).

[1] *Corresponding Author:*
T.W. (Taylor) Stone
t.w.stone@tudelft.nl

1 ENGINEERING ETHICS AS RESPONSIBLE INNOVATION

Over the past decade, the concept of Responsible Innovation (RI) has emerged as a guiding framework for technological innovation in the European context. At the most basic level, RI "is the on-going process of aligning research and innovation to the values, needs and expectations of society" [1]. Rejecting the position that technologies are value-neutral, RI maintains that moral values can be embedded in technologies, and that moral deliberation should be a fundamental element at all levels of technological research, development, and governance. RI can, as a concept, be understood via procedural (process-focused) and substantive (product-focused) dimensions [2]. As a procedural tool, RI explicates a process of innovation that meets identified societal norms (e.g., transparency, accountability, public participation), or that adheres to ethical principles such as non-maleficence and beneficence [3,4]. As a substantive notion, RI focuses on the outcomes of innovation, ensuring said processes result in artefacts or systems that positively foster identified moral values (e.g., safety, sustainability, privacy) [5].

While there are a variety of methodologies that fall under the umbrella of RI, two key unifying features can be highlighted. First, as an ethical approach to innovation, it is explicitly forward-looking. Rather than focusing on retrospective questions of responsibility and blame – a common approach to teaching engineering ethics – it asks how to develop technologies with, and for, society. Second, moral values are situated as fundamental considerations for engineering, design, and associated policy-making. Thus, the adherence to, or incorporation of, identified relevant values is situated as a "supra-functional" design requirement [6].

2 TEACHING RESPONSIBLE INNOVATION: TWO PROBLEMS

Given the constructive orientation of RI, the multiplicity of frameworks that have emerged to operationalize RI in different contexts, and the recognition of RI by governments, companies, and funding agencies, it comes as no surprise that it has also been also been incorporated into education. The concepts and tools of RI have become a fundamental component of engineering ethics training in various institutions, and ethics of technology courses more generally. However when teaching RI, and asking students to utilize different approaches as a research method, we typically rely on two intertwined (epistemic) assumptions:

- That we have a good sense of what the consequences of an intervention or innovation will be; and,
- That we know the meaning of a specific value in a specific context

Put otherwise, an assignment that asks students to "design X for value Y" implies we know both what the effects of (potential) innovations to X are, and have a clear and stable definition of Y. However, these assumptions should *not* be blindly accepted, especially when dealing with radical or disruptive technologies, which by definition have the potential to transform, in unforeseen ways, the values we take for granted. Accepting that these

assumptions should be addressed, we can therefore identify two interrelated challenges for teaching RI:

The Positivist Problem

The first problem calls into question the predictability of innovations and inventions, and the assumed linear relation between the design and use contexts of a technology. Based on theoretical and historical evidence, it has been argued that designer's intentions do not necessarily correspond with users practices. Rather, there is no essential or stable interpretation of a technology, but different uses that can emerge in different contexts [7,8]. Acknowledging the limits of foresight, RI literature rejects prediction as a goal, instead endorsing the importance of *anticipation*. Broadly put, anticipation is understood as an exploratory stance prompting "what if" questions, towards considerations of what is known, what is likely, and what is possible [3]. However, this opens up important pedagogical questions: what does *good* anticipation look like? How do we avoid excessively optimistic or pessimistic forms of anticipation? And, what activities give students the opportunity to develop this competency?

The Empty Signifier Problem

The second problem concerns the knowledge required to cultivate anticipation as a competency within engineering ethics education. Specifically, the goal of aligning innovation with societal values requires a nuanced understanding of what we *mean* by said values. However, this is not always clear. Values such as privacy or sustainability are so commonplace that we rarely question their origins, specificities, or legitimacy. Further, their (unquestioned) connotations can be co-opted to defend or reject a technology out-of-hand. This can result in superficial, ahistorical, and acontextual definitions – both in how such values are taught, as well as how students operationalize values in their assignments. While recent scholarship has drawn attention to this issue and proposed new mechanisms for addressing these deficiencies at a theoretical level, there is still the question of how to translate this rich (and evolving) debate into concrete teaching exercises. How can we structure exercises and assignments so that students move beyond a superficial identification of common values, towards acquiring a nuanced understanding of their import and meaning? How do we foster critical and reflective research into key values (*why does it matter, for whom, and how has it come to matter so much to us*)? And, how can students develop a sensitivity to the co-opting of values in arguments for or against a certain technology?

3 TEACHING RESPONSIBLE INNOVATION: FOSTERING 'HISTORICALLY INFORMED ANTICIPATION' AS A COMPETENCY FOR ENGINEERS

The *positivist* and *empty signifier* problems pose important epistemic challenges to the teaching (and practice) of RI, forcing us to question the limits of our knowledge. Yet, they need not be seen *only* as problems. Our proposal is that the above two challenges can be re-

framed as an opportunity to develop and refine RI teaching within engineering ethics, so that it confronts these issues head-on. This can be addressed, we propose, by focusing on a pre-condition for the successful application of RI theory. Anticipation can be situated as a procedural tool, and a benchmark for product development and associated policies. However, we can also position anticipation as a competency that should be explicitly fostered in the training of future engineers and designers, towards the goal of developing the knowledge and traits required for RI. Understood in this sense, we focus on anticipation as historically informed and as requiring the cultivation of a set of interrelated intellectual virtues – outlined below. Importantly, this competency is not bound by discrete knowledge and finite skills. Rather, it is about fostering a critical awareness of context and an attunement to the moral issues at stake therein. Thus, the virtues sketched below are about *how* to do RI, not *what* to do.

Moral sensitivity

We understand moral sensitivity as an overarching and fundamental requirement – for RI, as well as social and professional responsibility generally – that constitutes an attunement to *why* and *for whom* certain aspects of a situation or choice are morally relevant. As a facet of historically informed anticipation, it requires an ability to not just identify obvious values at stake in the design and introduction of a given technology, but also a sensitivity to a) the meaning of those values; b) the possibility of implicit biases obscuring the meaning of relevant values and the voices of relevant stakeholders; c) the interplay between identified values and the innovations that may affect the relevance we attach to them; and d) the presence of other (less obvious) values that may be at play too. Thus, moral sensitivity requires an attentiveness to the contextual meaning of values (e.g., the history of a certain value conflict).

Epistemic humility

Abandoning prediction in favour of anticipation requires that we also strive to engender a prudential outlook regarding possible futures. The social, environmental, and economic ramifications of emerging technologies are becoming increasingly complex and far-reaching. RI therefore requires a recognition of the limits of our knowledge about a technology, including the values it presumably instantiates and how innovations might affect different stakeholders, towards coping with the unforeseen consequences of failures *and* successes. The cultivation of a reasoned and critical approach can provide a middle-way between overly optimistic or pessimistic perspectives on new innovations [9].

Moral imagination

An anticipatory approach to technology development requires an ability to creatively explore the relationship between moral values and technologies, and to envision novel solutions to an identified (moral) problem. This should still be grounded in moral sensitivity and epistemic humility – in an attunement to ethical issues and some tempering of

uncritical techno-optimism. Yet, a hopeful approach to problem solving is essential for RI [9]. Coursework should therefore foster an open and exploratory outlook. This can be rooted in an acknowledgement of professional responsibility, while giving space to explore what sort of future we want, and why.

We believe that developing courses and exercises that explicitly aim to cultivate this set of intellectual virtues will inculcate the competency of *historically informed anticipation*. It will foster an appreciation of the limits of foresight, a critical perspective on the potential consequences of innovations and inventions, and a robust sense of values as socially, historically, and contextually contoured.

3.1 Historically informed anticipation in practice

Towards this goal, we propose a novel teaching methodology for engineering ethics, termed a *value-genealogy of technology*, focused explicitly on developing the competency of historically informed anticipation (Table 1). It is an open and explorative exercise, during which students will engage with a specific value-technology relation over the entire course – analysing how a value has shaped a technology and associated societal perceptions and policies, and vice versa. After an introduction to the main tenets of RI via readings and lectures, student groups will spend several weeks undertaking discussions and formative assignments related to their case. The exercises are intended to allow for an exploration of the historical relationship between the technology and the value, an analysis of meta-narratives shaping policies and public perceptions, and a critique of contemporary discourse.

Table 1. Overview of value-genealogy of technology assignment

Exercise	Objective	Output	Target Virtue
Writing 1: reflection	Reflection on initial impressions of the technology and value(s) to be explored	Short writing assignment	Moral sensitivity
Starter-kit analysis	Gain foundational knowledge via project description and initial resources provided by instructor	Identification of key values and (potential) conflicts	Moral sensitivity
Creating a genealogy	Critical analysis of historical texts, contemporary texts, and media	Mapping exercises of technology-value relations, conflicts, and predictions (past and present)	Moral sensitivity Epistemic humility
Presentation: (un)informed anticipation	Present results of genealogy via selecting two extreme perspectives (utopian and dystopian), and critically reflecting on the assumptions, biases, and context of those perspectives	Class presentation	Epistemic humility Moral imagination
Writing 2: informed anticipation	Analysis of initial impression, genealogy, and presentation; positing a future-oriented approach to case in line with RI	Essay	Epistemic humility Moral imagination

To help clarify and concretize this anticipatory exercise, the below box offers an overview of one value-technology case study that can be used in the exercise: the relationship between public order and (smart) lighting, resulting in a value conflict between safety, privacy, and surveillance. This text will constitute the introduction to the students' "starter kit," which will include readings on the social history of nighttime lighting, the ethics of smart cities, literature from companies advocating for smart lighting, and selected newspaper articles.

The Streetlights are Watching You: Values and Smart Lighting

Emerging "smart city" trends are spurring a new generation of streetlights, with lampposts being fitted with sensors, cameras, and a host of other novel technologies aimed at monitoring and data collection. While these innovations may offer improvements to efficiency and safety, they raise concerns about privacy, surveillance, and power dynamics. More fundamentally, such smart systems seemingly extend the technical functions of streetlights. No longer simply providing illumination, they actively monitor their environment and those who inhabit it, creating a vast network of nodes throughout our public spaces. Combined, the novel functions and capabilities of smart streetlights arguably create a new terrain of moral concerns. From such a vantage point, this technology acts as a socially disruptive force, profoundly altering the public spaces of cities and those who inhabit it. This has created a divide between the companies and cities championing the benefits of smart systems (the Utopians), and those who critique such technologies as socially and politically unjust (the Dystopians). What would a historically informed anticipatory intervention into this debate look like?

The history of nighttime lighting offers a nuanced perspective. Without denigrating contemporary concerns, we can find evidence that these seemingly novel issues represent a continuity with the values fundamental to the very notion of public lighting. Debates over social order at night – and the resultant tension between safety, privacy, and surveillance – have been a recurring theme for centuries. Streetlights have long been utilized as a form of policing and perceived as a symbol of authority, creating ongoing tensions between control and liberation in urban nightscapes. While offering significant improvements in accuracy, smart streetlights arguably embody a continuity of values – and value tensions – that can be traced back to the origins of public lighting in the 17th-18th centuries.

It seems that contemporary innovations represent new means of realizing these long-held goals, just as resistance to them offers fresh versions of protest and critique. But, do the possibilities of smart lighting technologies warrant a shifted perspective? Are these values (and value tensions) static, or do new innovations force us to re-think the meaning of notions like "surveillance in the public sphere"? How have perceptions of these values evolved with new lighting technologies, as well as social changes? Which stakeholders have a voice in (past and present) narratives about the technology? And, how are groups on both sides of this argument co-opting long-held ideas and associations (e.g., "lighting equals safety") to support their goals?

4 CONCLUSION

This act of developing a historically informed anticipation regarding a value-technology dynamic will – we propose – help students understand the context-specific meaning of a value, in turn allowing for a nuanced perspective on the potential social and environmental impacts of future innovations. While we will ask students to take a critical stance, as well as propose a constructive path forward, we are not asking for a final or definite "answer" to

the question “will innovation X count as an instance of RI?” Rather, we hope for a critical and reflective exploration of the mutual co-shaping of the value(s) and technology. We believe such an assignment will leave students – as engineers in training – better prepared to appreciate both the complexity of value-technology relations, and assist in habituating the virtues of moral sensitivity, (epistemic) humility, and moral imagination.

REFERENCES

[1] European Union (2014), Rome Declaration on Responsible Research and Innovation, accessed 15 March 2020 [https://ec.europa.eu/digital-single-market/en/news/rome-declaration-responsible-research-and-innovation-europe]

[2] von Schomberg, R., (2013), A Vision of Responsible Research and Innovation, *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, R. Owen, J. Bessant, M. Heintz (Eds.), John Wiley & Sons, Ltd., West Sussex, pp. 51-74.

[3] Stilgoe, J., Owen, R., Macnaghten, P. (2013), Developing a framework for responsible innovation, *Research Policy*, Vol.42, pp. 1568-1580.

[4] van de Poel, I., (2016), An Ethical Framework for Evaluating Experimental Technology, *Science and Engineering* Ethics, Vol. 22, No. 3, pp. 667-686.

[5] van den Hoven, J. (2013), Value Sensitive Design and Responsible Innovation, *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, R. Owen, J. Bessant, M. Heintz (Eds.), John Wiley & Sons, Ltd., West Sussex, pp. 75-83.

[6] van den Hoven, J. (2017), The Design Turn in Applied Ethics. *Designing in Ethics*, J. van den Hoven, S. Miller, T. Pogge (Eds.), Cambridge University Press, Cambridge, pp. 11-31.

[7] Albrechtslund, A. (2007), Ethics and technology design, *Ethics and Information Technology*, Vol. 9, pp. 3-72.

[8] Ihde, D. (2008), The Designer Fallacy and Technological Imagination. *Philosophy and Design: From Engineering to Architecture*, P. E. Vermaas, P. Kroes, A. Light, S. Moore (Eds.), Springer, Dordrecht, pp. 51-59.

[9] Vallor, S. (2016), Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting, Oxford University Press, Oxford, pp. 125-127.