Jun 23rd, 9:00 AM - Jun 28th, 5:00 PM

Navigating ethics-informed methods at the intersection of design and philosophy of technology

Deger Ozkaramanli
Delft University of Technology

Merlijn Smits
Saxion University of Applied Sciences

Maaike Harbers
Rotterdam University of Applied Sciences

Gabriele Ferri
Eindhoven University of Technology

Michael Nagenborg
University of Twente

See next page for additional authors

Citation

Ozkaramanli, D., Smits, M., Harbers, M., Ferri, G., Nagenborg, M., and van de Poel, I. (2024) Navigating ethics-informed methods at the intersection of design and philosophy of technology, in Gray, C., Ciliotta Chehade, E., Hekkert, P., Forlano, L., Ciuccarelli, P., Lloyd, P. (eds.), *DRS2024: Boston*, 23–28 June, Boston, USA. https://doi.org/10.21606/drs.2024.868

DRS BOSTON 20 DRS 24 JUNE 24 — 28

Navigating ethics-informed methods at the intersection of design and philosophy of technology

Deger Ozkaramanli[a,*], Merlijn Smits[b], Maaike Harbers[c], Gabriele Ferri[d], Michael Nagenborg[e], Ibo van de Poel[a]

[a]Delft University of Technology, the Netherlands
[b]Saxion University of Applied Sciences, the Netherlands
[c]Rotterdam University of Applied Sciences, the Netherlands
[d]Eindhoven University of Technology, the Netherlands
[e]University of Twente, the Netherlands

*Corresponding author e-mail: d.ozkaramanli@tudelft.nl

doi.org/10.21606/drs.2024.868

Abstract: The idea that technologies influence society—both positively and negatively—is not new. This is mainly the terrain of the philosophy and the ethics of technology research. Similarly, design research aims to help create new technologies in line with individual, social, and societal needs and values. Against this backdrop, it seems essential to expose relations between design and philosophy of technology research, particularly from a methodological perspective. The main goal of this paper is to suggest a preliminary overview of methods and approaches that can inspire and inform interdisciplinary collaboration and, with that, systematic engagement with ethics in design processes. Through interdisciplinary exchange, we propose a preliminary typology of ethics-informed methods and approaches based on two main dimensions, namely theory-grounded approaches to theoretically-flexible techniques and assessment to accompaniment. This mapping intends to help navigate the ethical qualities of selected methods from both disciplines, and it aims to create a platform for fruitful interdisciplinary conversations.

Keywords: design ethics; design methods; philosophy of technology; interdisciplinary collaboration

1. Introduction

People interact with numerous technologies every day, ranging from physical and digital products to services, systems, and spaces. These technologies serve utilitarian functions, but they also redefine or reinforce certain moral values and social practices. A public bench may prevent rough-sleeping, a children's playground may encourage social inclusion, and a coffee

mug may evoke tranquility. The idea that technologies are not neutral is not new (e.g., Winner, 1980; Verbeek, 2011). Most recently, philosophy of technology researchers have been grappling with the idea of techno-moral change, i.e., technologies change our norms and values, making it challenging to think of ethical frameworks as static concepts (e.g., Danaher& Sætra, 2023). This type of research is mainly the terrain of the philosophy and the ethics of technology, which offers concepts, language, and definitions to help understand and articulate *what technologies are* and to critically analyze *what technologies do* in society (de Vries, 2017).

Design researchers have also been interested in exploring the behavioral, cultural, social, and environmental effects of technologies. This interest helps create 'better' technologies that align with individual or societal needs and has given rise to various design methods and approaches, such as Social Implication Design (e.g., Tromp & Hekkert, 2018), Dilemma-Driven Design (Ozkaramanli, Desmet, & Ozcan, 2016), Participatory Design (e.g., Bjögvinsson, Ehn, & Hillgren, 2012), Critical Design (Dunne, 1999) and Speculative Design (Dunne & Raby, 2013). In this terrain, however, the ethics of these methods and approaches has not yet been explicitly discussed. For instance, the collection and categorization of design methods (e.g., IDEO DesignKit or Delft Design Guide by van Boeijen, Daalhuizen, & Zijlstra, 2020) rarely address the ethical qualities of the methods that they describe. Chivukula et al. (2021) analyzed 63 ethics- and values-focused design methods using content analysis. Although they found evidence for ethical theories such as deontological, consequentialist, virtue, pragmatist, and care ethics being explicitly mentioned, they also pointed out that method designers rarely indicate the ethical motivation behind their methods. Arguably, these motivations need not always be linked to a specific ethical theory. For instance, involving people as active participants in design processes (e.g., Participatory Design), changing behavior (e.g., Social Implication Design), or materializing critique (e.g., Critical Design) may be considered subtle ethical qualities. These qualities point to a need to foreground the ideological dimensions of design methods to facilitate interdisciplinary discussion.

As a result, the main goal of this paper is to sketch the landscape of methods and approaches to stimulate interdisciplinary collaboration between design and philosophy of technology researchers. The core assumption here is that methods guide ethical action, while not being the only factor that contributes to ethical behavior of designers. We build on the work by Steen (2015) that focused on design as a social process and compared Participatory Design, Human-Centered Design and Co-design using Virtue Ethics, Ethics of Alterity, and Pragmatist Ethics, respectively. This work revealed the hidden ethical qualities of these approaches when they are implemented in actual design practices. In addition, our work is inspired by the recent comprehensive review of ethics-focused design methods by Chivukula et al. (2021) and the work of Gray et al. (2023) to embed ethics in design practices. Differently from these works, our goal is not to form a comprehensive overview of ethics-focused methods, but to closely examine the main connections among selected methods and approaches to reveal interdisciplinary similarities and differences. With this, our main question

is: What are the main dimensions along which design methods and the philosophy of technology approaches relate to each other?

This paper is informed by conversations with eight experts (six academics and two practitioners) working at the intersection of philosophy of technology research, design research, and innovation practices where design plays a role. In addition, we conducted subsequent literature research based on the sources and methods suggested by these experts. To clarify, the terms 'method' and 'approach' are widely used in design literature, and it is common to read inconsistent usage of these terms. In this paper, a 'method' refers to mental tools that aims to provide structure and support in dealing with complex and complicated problems in varying projects, contexts, and environments (Badke-Schaub, Daalhuizen, & Roozenburg, 2011), and an approach is the more open-ended term that refers to the mindset or belief-system with which a method is executed. Finally, borrowing from Sanders, Brandt, and Binder (2010), we define tools as tangible components that are used in project activities, such as cards, templates, or infographics; and technique refers to a description of how one or more tools can be implemented in project activities (e.g., card sorting or scenario creation).

In what follows, we highlight methods and approaches in the philosophy of technology and design research, respectively. Next, we propose a preliminary typology for a relational perspective on these methods and approaches, which is intended as a conversation tool to inform and inspire interdisciplinary collaboration. Finally, we discuss implications for interdisciplinary research.

2. Philosophy of technology approaches

In this section, we summarize selected approaches from philosophy of technology that illustrate how to engage with philosophical and ethical reasoning about existing and emerging technologies. These approaches help interpret the role of science and technology in society beyond everyday interpretations or dominant discourses. Perhaps most visible are questions regarding the assessment of the social and ethical impact of specific technologies, which have historically been addressed by *Technology Assessment* (TA). Next, we turn to what we call *anticipating-sensing* approaches that can inform design processes (vs. assess outcomes). Finally, we discuss *accompaniment* approaches that can guide design processes from within the process.

TA approaches originated in the 1960s in the United States of America, and they help monitor the social, legal, economic, and ethical impact of technological developments on society to inform policymakers on alternative policies (Banta, 2009). It is an extremely broad and constantly evolving field with sub-fields such as Health Technology Assessment (Banta, 2003) that provides input for decision-makers on health technology options. Traditional TA approaches rely on expert knowledge, quantifiable risks, and are conducted by institutions outside of technology development (e.g., U.S. Office of Technology Assessment, research institutions). This means that TA does not aim to directly influence or broaden design processes,

but to monitor their boundaries. Practically, TA evaluates technological impact based on themes such as autonomy, dignity, benevolence, accessibility, justice, transparency, and social interaction. Compared to traditional TA, Constructive Technology Assessment (CTA) aims to include a larger variety of actors (e.g., social, technological actors) in predicting impact and emphasizes the importance of assessing and addressing the social implications of technologies *during* their development (Schot & Rip, 1997). Building on CTA, Ethical CTA (eCTA) (Kiran, Oudshoorn, & Verbeek, 2015) argues for a situated (vs. universalized) and nuanced (vs. focused purely on adverse effects) form of assessment that considers the dynamic nature of moral values and ethical principles. With eCTA, assessment becomes more co-creative, more mindful of individual and cultural differences, and more in line with a co-evolutionary approach to technological impact (i.e., the idea that society co-evolves with technology (Geels, 2005). These characteristics (i.e., co-creative, situated, and nuanced assessment) echo usability and experience evaluation methods in design research, although the latter mostly focuses on individual and not on the societal context of technologies.

Anticipating-sensing approaches, such as Technological Mediation (Verbeek, 2011) and Technomoral Change Scenarios (Swierstra, Stemerding, & Boenink, 2009), employ anticipation exercises to pinpoint or 'sense' key concerns and values around emerging technologies or technological ideas (e.g., Google Glass in Kudina & Verbeek, 2019). In that sense, and unlike traditional TA, they aim to directly influence or broaden design processes. For instance, Theory of Technological Mediation deems technologies as *mediators* of human-world relations, recognizing that everyday experiences are fundamentally shaped by technologies that we use (e.g., an ultrasound examination of a fetus shapes perceptions and expectations from parenthood) (Verbeek, 2011). This helps technology developers to analyze, anticipate, and experiment with the relations between humans and products, and with that, how technologies impact human experience, behavior, and values. Product-Impact Tool implements and builds on mediation analysis to facilitate ethical engagement and discussion about the impact of technologies on people, society, and the environment in a practical and engaging manner (Dorrestijn, 2020). In addition, Technomoral Change Scenarios explore the emotions and controversies that new and emerging scientific and technological developments provoke through stimulating moral imagination using narratives (see the case of the obesity pill, Swierstra, Stemerding, & Boenink, 2009). In summary, anticipating-sensing approaches reveal present-day concerns and values so that they can be inserted into design processes, although examples of designer-philosopher collaborations using these approaches are rare.

Anticipation is viewed critically by the Technology-as-Social-Experiment approach, which focuses on regulating the innovation process (vs. speculating about its outcomes) (van de Poel, 2011; 2013). It acknowledges the radical uncertainties and potential hazards of new technologies (e.g., nanotechnology, biotechnology) and aims to deal with them through an adaptive learning process similar to scientific and medical experiments. To support a responsible introduction of new technologies, this approach offers four general ethical principles, namely nonmaleficence, beneficence, respect for autonomy, and justice. Some of these principles refer to the need to be able to adjust a technological design after a technology has entered

society. This may, for example, be affected by following design strategies like adaptability and flexibility (van de Poel, 2021; van der Weij et al., 2023). It also implies a view of design in which the design process does not end once a technology has entered society, but in which technologies can be redesigned based on new insights on ethical implications that became clear after a technology has been introduced into society.

Finally, *accompaniment approaches* include Guidance Ethics (Verbeek & Tijink, 2020) which advocates for a bottom-up and positive approach that provides concrete actions for those involved in technology development. Guidance ethics involves a workshop in which various stakeholders (ranging from citizens to technology developers and managers) interact with each other and brainstorm about the positive and negative outcomes of technology. This workshop can be done in the conceptualization phase of a technology, but also during development or implementation. Its outcomes relate to concrete action opportunities for technology (hardware and software), technology-users (behavior), and context (e.g., education/management/policy) (e.g., Siebelink et al., 2024).

In summary, the spectrum from assessment to anticipating-sensing and finally to accompaniment represents to what extent an ethical evaluation of a technology or a technological idea happens from outside or from within the design process: Assessment approaches *assess* design outcomes from outside the process, anticipating-sensing approaches *inform* design processes, and accompaniment approaches directly *guide* design processes from within the process.

3. Design methods and approaches

In this section, we summarize selected approaches from design research that possess implicit ethical qualities. Designing is charged with ethical questions and dilemmas. Whitbeck (1996) famously drew an analogy between design problems and moral problems, arguing that philosophers can learn from design reasoning to respond to moral problems. This analogy highlights that design reasoning can aid ethical reasoning through dealing with uncertainty and being open to the dynamic character of problem situations and pursuing viable solutions instead of the 'best' solution. However, apart from Whitbeck's (1996) and the work of a few other scholars (e.g., Gray et al., 2023), how design methods may implicitly stimulate or hinder ethical engagement remains unexplored. Here, we aim to unpack the ethics of several design methods that demonstrate ethical qualities (e.g., ethical intentions, value judgments, opportunities for ethical engagement). Drawing from Vermaas et al. (2015), we present our analysis in three main categories: designer-driven approaches, stakeholder-driven approaches, and those driven by specific ethical principles.

Designer-driven approaches emphasize the agency of the designer in steering a collaborative project. An example of this is the Vision in Product Design (ViP) (Hekkert & van Dijk, 2011) which emphasizes the freedom, authenticity, and responsibility of the designer as a societal actor. ViP invites engaging in design projects through forming a statement or a vision that

balances people's needs with the designer's interpretation of structural factors that influence society (e.g., technological progress, economic factors, psychological factors, socio-cultural developments). The Social Implication Design (SID) method (Tromp & Hekkert, 2018) builds on ViP to support designers in reasoning from a social issue to a design proposal through focusing on social dilemmas. In this way, it exploits the implicit yet inevitable role of design in changing human behavior in socially desired directions (e.g., social cohesion, healthy living). Similarly, Dilemma-Driven Design (Ozkaramanli, Desmet, & Ozcan, 2016) considers personal dilemmas as valuable starting points for understanding people and conceiving innovative design ideas. This brings a human-centered focus to technological discussions by explicating the mixed emotions and conflicting concerns that people may experience in response to new technologies.

A distinct group of designer-driven approaches focuses on revealing radically different technological possibilities specifically to critique them. For instance, Critical Design (Dunne, 1999) produces provocative artefacts that challenge consumerist desires and societal norms. Critical Design is not concerned with immediate utility but acts as 'food for thought', highlighting possible implications of present-day technologies, often merging insights from ethics, philosophy, political science, and more. Speculative Design (Dunne & Raby, 2013) is closely related to Critical Design, with which it shares a critical ethos, but focuses on parallel and otherwise-possible realities rather than futures. It prompts 'what if' questions that problematize taken-for-granted assumptions. Speculative Design serves as the starting point for two related methods, namely, Design Fiction (Sterling, 2005; Bleeker et al., 2022) and Adversarial Design (DiSalvo, 2012). While they share some roots, their attitude towards commercial design practices varies: Design Fiction is an industry-ready way of materializing possible future products into 'diegetic prototypes' to probe their possible reception and ethical ramification, while Adversarial Design – as its name suggests – is a more radical way to design 'against' expectations to challenge the status quo. While all these approaches differ in their stylistic and critical components, they all share the aim of shining a light on seductive but problematic design trends. Because of this, they relate to anticipating-sensing approaches discussed in the previous section, but with an added experiential layer that makes complex ethical ideas easier to understand.

Stakeholder-driven approaches emphasize distributed responsibility of actors involved in design and innovation projects. They are generally not informed by specific ethical theories, and the outcomes are open-ended and informed largely by the positions and views of the involved stakeholders. An example of a stakeholder-driven approach is Participatory Design (e.g., Björgvinsson, Ehn, & Hillgren, 2012), which emerged in the 1970s in Scandinavia and is rooted in democratizing workplaces (e.g., how to manage division of labor, how to implement new production methods and tools). Participatory Design gives voice to those who are most influenced by technologies in the development process of new technologies, which "reflects the then-controversial political conviction that controversy rather than consensus should be expected around an emerging object of design" (Björgvinsson, Ehn, & Hillgren, 2012, pp. 103). There are many variations of participatory design, though some may have

lost the initial political spirit (Bannon, Bardzell, & Bødker, 2018). Value Sensitive Design (Friedman, 1996; Friedman & Hendry, 2019) is another example of a stakeholder-driven approach that aims to embed values in technology design through technical, conceptual, and empirical research. Similarly, by studying the lived experiences of stakeholders in their daily context through action research and living labs, the Values that Matter method focuses on analyzing and integrating the dynamic nature of values into technology design and development (Smits et al., 2022).

Approaches driven by ethical principles center around the significance of values deemed beneficial for society at large. Discussions on risks and harms have recently gained traction, especially in the realm of Artificial Intelligence (AI) and Large Language Models (Ferri & Gloerich, 2023) with recommendations for a human rights-oriented approach to AI regulation and its effects on marginalized communities (Bender et al., 2021; Prabhakaran et al., 2022; Fjeld et al., 2020). In addition, Privacy-by-Design (Hustinx, 2010) and Ethics-by-Design (European Commission, 2021) highlight the importance of minimizing data collection and respecting privacy as a fundamental value in the development and/or use of AI-based systems.

More politically-charged examples include Feminist Human-Computer Interaction (Bardzell & Bardzell, 2011), which challenges and reimagines traditional methods by highlighting the gendered nuances of technology and underscoring the importance of inclusion, diversity, and agency. Other methods, like those highlighted by Toombs et al. (2016), are inspired by care ethics (de la Bellacasa, 2011). They highlight relational ties, responsibilities, and the need to address specific individual concerns, urging for designs that embrace compassion and empathy. The Design Justice framework (Costanza-Chock, 2020) critically examines how design can disproportionately distribute burdens, rewards, and risks among different societal groups. Grounded in intersectionality, Design Justice advocates for the evaluation of multiple socio-economic and historical factors when determining the allocation of benefits and harms. This ensures that justice remains at the core of design decisions. In this context, the principle of justice is foundational to Escobar's (2018) concept of pluriversal "autonomous design". This approach promotes harmony with nature, prioritizes community collaboration over commercial interests, and shifts away from perpetual modernization. It engages with environmental, experiential, and political concerns, emphasizing the entanglement of all beings. The essence of these ethical-principle-driven approaches is not merely to infuse designs with ethical considerations but to anchor them as the very foundation upon which designs are conceived and realized.

In summary, the ethical qualities of the aforementioned design methods are multiple: They demonstrate ethical intentions (e.g., participation, provocation, protecting privacy); they involve value judgments as part of procedural decisions; and they add an experiential dimension to abstract discussions around values and ethics through experientable artefacts and with that, determine and steer these discussions.

4. Towards a typology

In this section, we propose a preliminary typology of methods, which serves as a living document and a conversation tool to inspire researchers from philosophy of technology and design to engage in interdisciplinary collaboration.

4.1 Mapping out the relations among methods

The **y-axis** (see Figure 1) (theory-grounded approaches to theoretically- flexible techniques) represents to what extent a method can be traced back to specific theories and whether it offers actionable insights on how to implement it in design processes. If we could trace the theoretical basis of a method in literature (e.g., agonistic democracy for Adversarial Design, DiSalvo 2012), we placed them high on the y-axis. If the theoretical basis of the method was unclear, but could be inferred from the literature, we placed them approximately in the middle part of the y-axis. If there was no direct theoretical basis, or when there were multiple or varying theoretical bases, to which we refer as theoretically-flexible-methods, we placed them low on the y-axis. In that case, we assumed that such methods are probably techniques (vs. methods).

The **x-axis** (assessment to accompaniment) represents the type of ethical engagement that is enabled by a method: assessment approaches monitor design from outside the process, anticipating-sensing approaches *inform* design; and accompaniment approaches *guide* design from within the process. This dimension also denotes a historical development in the ethics of technology as participatory approaches are increasingly popular forms of ethical engagement alongside constantly evolving TA approaches.

As a result, we propose to cluster methods and approaches that we described in the previous two sections, along two main dimensions and four quadrants. We explain each quadrant in the following paragraphs.

4.2 Theory-grounded, assessment approaches

We place the traditional Technology Assessment (TA) methods in this quadrant due to their focus on forecasting technological impact. Traditional TA also includes the legal aspects of technological development, for which legal scholars are involved to ensure compliance with pre-determined rules and regulations (e.g., product safety, privacy) and to ensure accountability. Building on traditional TA, methods such as CTA and eCTA involve a larger variety of stakeholders in the assessment protocols and aim for more nuanced and situated analysis (Schot & Rip, 1997; Kiran, Oudshoorn, & Verbeek, 2015). These TA methods add an explicit ethical dimension to traditional TA and increasingly focus on innovation policies to stimulate desirable technological developments (e.g., Grunwald, 2011).

Critique-oriented design approaches, such as Critical Design (Dunne, 1999; Dunne & Raby, 2013) and Adversarial Design (DiSalvo, 2012), also fall in this space. Although they are not 'assessment' approaches in the typical sense, they have a 'monitoring' function as they critically reflect on ongoing scientific and technological developments to inform reflexivity in the

design discipline. These approaches are grounded in philosophy and political theory, although Critical Design has not been explicit about its theoretical grounding. At the border of this space are the anticipating-sensing approaches, which are Technological Mediation (Verbeek, 2011) and Technomoral Change Scenarios (Swierstra, Stemerding, & Boenink, 2009), which originate in philosophy of technology research.

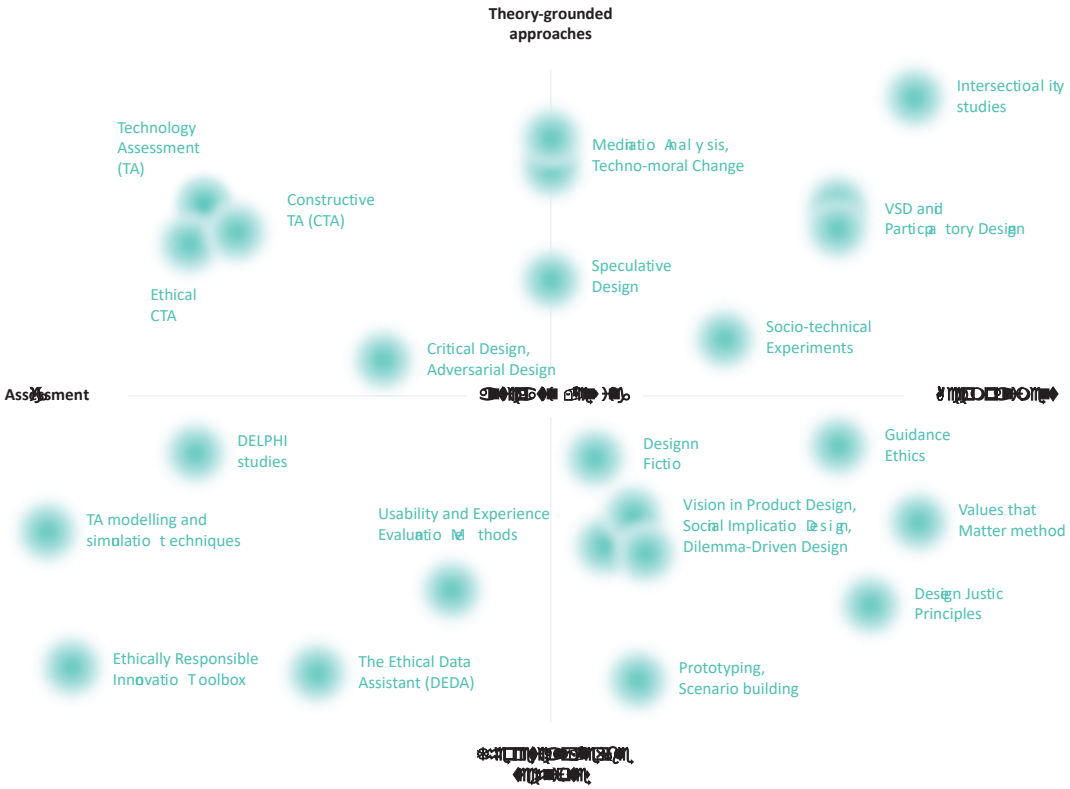

Figure 1 The interdisciplinary conversation space created by two main dimensions: theory-grounded approaches to theoretically-flexible techniques and assessment to accompaniment; and the preliminary mapping of ethics-informed methods and approaches discussed in this paper. Note that the location of the methods is not absolute but an estimation relative to the location of other methods in the same quadrant.

4.3 Theory-grounded accompaniment approaches

We argue that approaches in this quadrant are, to differing extents, grounded in theory. For instance, Socio-Technical Experimentation does not assume a specific ethical theory, but it builds on principles from research ethics and biomedical ethics and is based on certain theoretical assumptions about the relation between technology and society. Value Sensitive Design emphasizes integration of conceptual, empirical, and technical explorations (Friedman & Hendry, 2019), and Participatory Design is loosely grounded in particular visions for democracy (Bannon, Bardzell, & Bødker, 2018). We placed Speculative Design on the cusp between accompaniment and assessment to indicate that its stated goal of asking provocative 'what if' questions could potentially serve both ends. As we move closer to the accompaniment end of the quadrant, the intended interweaving of approaches with design practices

increase. At the extreme end of this quadrant is an illustrative example of a theory-grounded approach, namely intersectionality studies. Other examples might include Capability Approach (Nussbaum, 2011) and Feminist Care Ethics (Toombs et al., 2016). The main idea here is that, committing to a specific theoretical lens guides decisions throughout the design process.

4.4 Theoretically-flexible accompaniment techniques

We populate this quadrant mostly with design methods and approaches that offer actionable guidance ('technique') that can accompany designers and other technology developers during their activities. This is because the task of interpreting that specific theory and transforming it to actionable design guidelines (e.g., Design Justice principles by Costanza-Chock, 2020) rest with designer-researchers, unless previous research and/or case studies offer guidance that can be appropriated. Guidance Ethics (Verbeek & Tijink, 2020) and Values that Matter (Smits et al., 2022) are developed specifically for this purpose.

Vision in Product Design (Hekkert & van Dijk, 2011), Social Implication Design (Tromp & Hekkert, 2018), and Dilemma-Driven Design (Ozkaramanli, Desmet, and Ozcan, 2016) offer opportunities for ethical engagement and guidance on how to deal with emerging value conflicts in design projects. Design Fiction (Sterling, 2005; Bleecker et al., 2022) focuses on visualizing possible worlds where specific artefacts exist, allowing designers to showcase and evaluate their ethical desirability. Although it shares origins with Speculative and Critical Design, it does not rely on critical theory and, therefore, we place it in the theoretically-flexible quadrant.

Theoretically-flexible techniques often include scenario creation, which has been implemented in various forms such as Techno-moral Scenarios (Swierstra, Stemerding, & Boenink, 2009), value scenarios (Friedman & Hendry, 2019), socio-technical scenarios (Rip & Kulve, 2008) or design fiction (Sterling, 2005; Bleecker et al., 2022). Another popular technique is prototyping activities that can be designed in a way that adds an experiential dimension to discussions around values and ethical principles, rendering the topic accessible to a wider audience. Prototypes are sometimes built in an intentionally provocative manner (i.e., provotypes, e.g., Boer & Donovan, 2012) to stimulate debate and elicit deeper emotions and value discussions among stakeholders.

4.5 Theoretically-flexible assessment techniques

Mirroring TA approaches, this quadrant includes concrete protocols and formats of assessment widely used across philosophy of technology and design. For instance, Delphi is a scientific method that helps to organize an expert discussion to generate insights on controversial topics that result from rapid technological and social change (e.g., Beiderbeck et al., 2021). In addition, a variety of computational modelling and simulation techniques are used to execute TA. Moreover, specific 'toolboxes' are developed in governance contexts, such as

the Responsible Research and Innovation Toolkit[1], Ethically Responsible Innovation Toolbox[2], the Ethical Data Assistant[3], to facilitate responsible practices. Finally, in design research, various usability and experience evaluation methods exist to test and redesign existing technologies in line with design requirements (e.g., accessibility, durability).

5. Discussion: What this preliminary typology does and does not do

The visual manifestation of a preliminary typology in Figure 1 has helped to examine various methods and approaches from design and the philosophy of technology research. The core assumption here is that methods guide ethical action, while not being the only factor that contributes to ethical behavior of designers. This preliminary comparison of methods reveals two main insights to guide interdisciplinary collaboration. First, design researchers are invited to explicate the theoretical grounding of the methods that they develop, and with that, to reflect upon how these methods guide value commitments and help or hinder ethical reflection. For instance, a scenario-based design process driven by feminist care ethics (Toombs et al., 2016) as a theoretical lens will be different than one that is driven by Capability Approach (Nussbaum, 2011) or one that is not grounded in theory.

Second, philosophy of technology researchers are invited to engage with design practices through making abstract theoretical ideas and principles concrete by co-developing tools and techniques to implement them in real-life projects. Our observation is that most philosophy approaches are characterized by specific theories or ethical principles, yet they do not always offer procedural knowledge (e.g., methods, techniques) to guide their implementation. Even those who do include guidelines (e.g., Verbeek, 2011; Swierstra, Stemerding, & Boenink, 2009) may remain vague and not sufficiently actionable in design research and practices. This is to be expected as philosophy researchers may not always have an in-depth understanding of design practices, and this, once again, underlines the importance of interdisciplinary collaboration. In addition, better understanding 'how technologies are created' may help developing a more nuanced understanding of design through a central concept to the discipline, i.e. its methods. In this way, they can focus on particular practices in design, such as participatory or critical; instead of focusing on design in a general sense.

The third benefit of this diagram is pragmatic: To guide interdisciplinary conversations by raising questions around where to position new methods or approaches, and more importantly, why. We argue that the vertical dimension is important here as it stimulates design researchers to develop or explicate definitions and theoretical commitments, which is common practice in philosophy research. Alternatively, this dimension stimulates philosophy of technology researchers to engage with method research and development to increase the uptake of philosophy of technology theories and approaches. Here, there is no one-to-one

[1] Responsible Research and Innovation Toolkit: https://rri-tools.eu/

[2] Impact Assessment Human Rights and aI Algorithms (Impact Assessment Mensenrechten en Algoritmes, IAMA, in Dutch): https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes

[3] The Ethical Data Assistant (De Ethische Data Assistant, DEDA, in Dutch): https://deda.dataschool.nl/

relationship between certain techniques and theoretical lenses (hence they are theoretically-flexible). This also means that not all techniques would serve equally well when implementing theory in practice. This may stimulate discussions on which technique may work best for a specific theoretical lens and why.

In addition to direct interaction with specific philosophical theories and approaches, we encourage method-developers in design research to actively engage with philosophical reasoning to better communicate the subtle ethical qualities that underpin their methods. For instance, Dilemma-Driven Design (Ozkaramanli, Desmet, & Ozcan, 2016) proposes three directions for dealing with dilemmas but does not engage in philosophical reflection on what these dimensions could mean in human-technology relations. One of these directions is to moderate dilemmas through suggesting behavioral priorities. This type of 'moderation' might be problematized through the question of paternalism (Gertz & Ozkaramanli, 2024). In fact, the question of paternalism applies not only to design methods that explicitly aim to change behavior, but also to hidden forms of paternalism in, for example, user-centered design (see Gertz & Ozkaramanli, 2024 for a critical discussion). As a result, this emerging typology may invite method researchers to be explicit and precise about the ethical qualities of their methods when thinking about where to place them on the diagram.

We argued for placing most stakeholder- and designer-driven approaches on the accompaniment end of the assessment-accompaniment spectrum. The challenge here might be to balance emergent values in stakeholder- or designer-driven approaches with ethical principles. This is because 'values' may be understood differently through different disciplinary lenses or professions. Moreover, it is tempting, for example, to give into the straightforward definition that values are 'what people value'. Yet, this runs the risk of letting values such as 'racism' or 'pure profitism' to guide design processes simply because some stakeholders may value them (hypothetically speaking). For this reason, it is important to balance emergent values with values that the ethics of technology endorses in a normative sense (e.g., social equality, fairness, safety).

Finally, we acknowledge the challenge that researchers and/or designers may not reflect deeply on the methods they use in their activities. At the same time, we recognize the need to critically engage design researchers and designers with the ethical dimensions of the methods they develop and/or use (Ozkaramanli & Nagenborg, 2024). Choosing a method is an ethical decision guided by the values of an individual, and in turn, methods may guide designers in incorporating values in their activities that are not necessarily their own. As a result, we see that the design process is influenced both by methodological choices and by the actions and values of individuals involved in the process. In fact, a core design skill might be to negotiate value conflicts and moral dilemmas that arise in such collaborations.

5.1 Limitations

The proposed preliminary typology is an initial mapping of selected methods and its usefulness demands empirical validation. Implementing it in future interdisciplinary projects is part

of our future research agenda. In addition, our typology does not capture the nuanced distinctions between various theoretical commitments. Although the spectrum spans from assessment to accompaniment, we intentionally did not delve into the intricate differences between theories and their substantive content. Instead, we provided a panoramic view to give a general impression of the relationships between the included approaches, inviting readers to explore more precise distinctions themselves. For instance, theoretical perspectives like critical theory and feminist theory might both appear similarly on the spectrum, yet their foundational premises and implications vary greatly. Misinterpreting this diagram could result in oversimplifying the complex world of design ethics, which is an outcome we want to avoid. While our typology helps in understanding these theories' functional dynamics in design contexts, there is a risk it could overshadow their content and the processes that they advocate or critique. Thus, while the figure serves as a conversation tool for positioning theories on the assessment-accompaniment continuum, we urge researchers and practitioners to view it as a general map, and then delve into each theory and method to apply them comprehensively.

6. Conclusion

The main goal of this paper was to identify meaningful dimensions along which methods and approaches from design and philosophy of technology research can be compared and discussed. For this, we proposed an emerging typology based on two main dimensions: (1) theory-grounded approaches to theoretically-flexible techniques; and (2) assessment to accompaniment. Using this typology, we discussed similarities and differences among a variety of methods and approaches. Ultimately, we aim for this preliminary typology to act as a conversation tool to facilitate interdisciplinary collaboration between design researchers and philosophy of technology researchers.

We recognize that this preliminary typology is limited by the expertise of the contributing authors, and thus, it is not a comprehensive overview. Here, we were able to concretely discuss a small number of methods that we are familiar with and map out concepts and questions that can stimulate an interdisciplinary conversation. In line with this limited scope, this typology is intended as a starting point for discussion, for example, to compare methods, to argue for or against their positioning on the diagram, to think of new methods that may fall in a specific quadrant, or to formulate new research questions or alternative dimensions. One might also think of other ways of mapping out all methods based on, for example, a temporal dimension (i.e., methods for idea generation to methods for technology implementation), scale (individual to societal), or stakeholder involvement (top-down to bottom-up). In other words, we suggest judging the added value of this emerging typology by the quality of the discussion that it generates and envision it to act as a conversation tool when designing interdisciplinary collaboration.

Acknowledgements: This research is supported by the Dutch Creative Industries (ClikcNL) and contributes to the research agenda on the Key Enabling Methodologies (KEMs) for Mission-Driven Innovation. We thank Olya Kudina, Tjerk Timan, and Nazli Cila

for their helpful comments on an earlier version of this paper. Our work has been partially funded by Delft Design for Values Institute. Ibo van de Poel's contribution is part of the project ValueChange (ERC, Horizon 2020: 788321) and Ethics of Socially Disruptive Technologies (NWO: 024.004.031).

7. References

Badke-Schaub, P., Daalhuizen, J., & Roozenburg, N. (2011). Towards a designer-centred methodology: Descriptive considerations and prescriptive reflections. In B. Herbert (Ed.), *The future of design methodology* (pp. 181-197). Springer.

Bannon, L., Bardzell, J., & Bødker, S. (2018). Reimagining participatory design. *Interactions*, *26*(1), 26-32. https://doi.org/10.1145/3177794

Banta D. (2003). The development of health technology assessment. *Health Policy (Amsterdam, Netherlands)*, *63*(2), 121–132. https://doi.org/10.1016/s0168-8510(02)00059-3

Banta, D. (2009). What is technology assessment?. *International Journal of Technology Assessment in Health Care*, *25*(S1), 7-9. https://doi:10.1017/S0266462309090333

Bardzell, S., & Bardzell, J. (2011). Towards a Feminist HCI Methodology: Social Science, Feminism, and HCI. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 675–684. https://doi.org/10.1145/1978942.1979041

Beiderbeck, D., Frevel, N., von der Gracht, H. A., Schmidt, S. L., & Schweitzer, V. M. (2021). Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. *MethodsX*, *8*, 101401. https://doi.org/10.1016/j.mex.2021.101401

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

Björgvinsson, E., Ehn, P., & Hillgren, P. A. (2012). Design things and design thinking: Contemporary participatory design challenges. *Design issues, 28*(3), 101-116. https://doi.org/10.1162/DESI_a_00165

Bleecker, J., Foster, N., Girardin, F., & Nova, N. (2022). The Manual of Design Fiction. Near Future Laboratory.

Boer, L., & Donovan, J. (2012, June). Provotypes for participatory innovation. In *Proceedings of the designing interactive systems conference* (pp. 388-397). https://doi.org/10.1145/2317956.2318014

Chivukula, S. S., Li, Z., Pivonka, A. C., Chen, J., & Gray, C. M. (2021). Surveying the landscape of ethics-focused design methods. arXiv [cs.HC]. https://doi.org/10.48550/arXiv.2102.08909

Costanza-Chock, S. (2020). Design justice: Community-led practices to build the worlds we need. MIT Press.

Danaher, J., & Sætra, H. S. (2023). Mechanisms of Techno-Moral Change: A Taxonomy and Overview. *Ethical Theory and Moral Practice*, 1-22. https://doi.org/10.1007/s10677-023-10397-x

De la Bellacasa, M. P. (2011). Matters of care in technoscience: Assembling neglected things. Social Studies of Science, 41(1), 85–106. https://doi.org/10.1177/0306312710380301

De Vries, M. J. (2017). Philosophy as critique. In P. J. Williams & K. Stables (Eds.), *Critique in Design and Technology Education* (pp. 15–30). Springer Nature.

DiSalvo, C. (2012). Adversarial design. MIT Press.

Dorrestijn, S. (2020). A Tool for the Impact and Ethics of Technology: The Case of Interactive Screens in Public Spaces. In H. Wiltse (Ed.) *Relating to Things: Design, Technology and the Artificial*, (pp. 151-172). Bloomsbury.

Dunne, A. (1999). Hertzian tales: Electronic products, aesthetic experience and critical design. RCA CRD Research Publications.

Dunne, A., & Raby, F. (2013). Speculative Everything: Design, Fiction, and Social Dreaming. MIT Press.

Escobar, A. (2018). Designs for the pluriverse: Radical interdependence, autonomy, and the making of worlds. Duke University Press.

European Commission. (2021). Ethics By Design and Ethics of Use Approaches for Artificial Intelligence. Report last accessed on 31 October 2023 at https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf

Ferri, G., & Gloerich, I. (2023). Risk and Harm: Unpacking Ideologies in the AI Discourse. Proceedings of the 5th International Conference on Conversational User Interfaces, 284–289. https://doi.org/10.1145/3571884.3603751

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Research Publication*, (2020-1). http://dx.doi.org/10.2139/ssrn.3518482

Friedman, B. (1996). Value-sensitive Design. Interactions, 3(6), 16–23. https://doi.org/10.1145/242485.242493

Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press.

Geels, F. W. (2005). Processes and patterns in transitions and system innovations: Refining the co-evolutionary multi-level perspective. *Technological forecasting and social change*, *72*(6), 681-696. https://doi.org/10.1016/j.techfore.2004.08.014

Gertz, N. & Ozkaramanli, D. (2024). Beauvoir versus Behavior Change: Introducing Existential Ethics to the Politics of Design. *Design Issues, 40*(2), in print.

Gray, C. M., Chivukula, S. S., Carlock, T. V., Li, Z., & Duane, J. N. (2023, July). Scaffolding Ethics-Focused Methods for Practice Resonance. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (pp. 2375-2391). https://doi.org/10.1145/3563657.3596111

Grunwald, A. (2011). Responsible innovation: bringing together technology assessment, applied ethics, and STS research. *Enterprise and Work Innovation Studies*, *31*, 10-9. http://hdl.handle.net/10362/7944

Hekkert, P. P. M., & Van Dijk, M. B. (2011). *Vision in design-A guidebook for innovators*. BIS publishers.

Hustinx, P. (2010). Privacy by design: delivering the promises. *Identity in the Information Society*, *3*(2), 253-255. https://doi.org/10.1007/s12394-010-0061-z

IDEO (2023). https://www.designkit.org/

Kiran, A. H., Oudshoorn, N., & Verbeek, P. P. (2015). Beyond checklists: toward an ethical-constructive technology assessment. *Journal of Responsible Innovation*, *2*(1), 5-19. https://doi.org/10.1080/23299460.2014.992769

Kudina, O., & Verbeek, P. P. (2019). Ethics from within: Google Glass, the Collingridge dilemma, and the mediated value of privacy. *Science, Technology, & Human Values*, *44*(2), 291-314. https://doi.org/10.1177/0162243918793711

Nussbaum, Martha. 2011. *Creating Capabilities; The Human Development Approach*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Ozkaramanli, D. & Nagenborg, M. (2024). Moral Engagement in Design: Five Considerations for Unpacking the Ethical Dimensions of Design Methods. *Design Issues, 40*(3), in print.

Ozkaramanli, D., Desmet, P. M., & Ozcan, E. (2016). Beyond resolving dilemmas: Three design directions for addressing intrapersonal concern conflicts. *Design Issues*, *32*(3), 78-91. https://www.jstor.org/stable/24916448

Prabhakaran, V., Mitchell, M., Gebru, T., & Gabriel, I. (2022). A Human Rights-Based Approach to Responsible AI (arXiv:2210.02667). arXiv. https://doi.org/10.48550/arXiv.2210.02667

Rip, A., & Kulve, H. T. (2008). Constructive technology assessment and socio-technical scenarios. In *Presenting futures* (pp. 49-70). Springer Netherlands.

Sanders, E. B. N., Brandt, E., & Binder, T. (2010). A framework for organizing the tools and techniques of participatory design. In *Proceedings of the 11th Biennial Participatory Design Conference* (pp. 195-198). https://doi.org/10.1145/1900441.1900476

Schot, J. W., & Rip, A. (1997). The past and future of constructive technology assessment. *Technological Forecasting and Social Change,* 54 , 251–268. https://doi.org/10.1016/S0040-1625(96)00180-1

Siebelink, N. M., van Dam, K. N., Lukkien, D. R., Boon, B., Smits, M., & van der Poel, A. (2024). Action Opportunities to Pursue Responsible Digital Care for People With Intellectual Disabilities: Qualitative Study. *JMIR Mental Health*, *11*, e48147. https://mental.jmir.org/2024/1/e48147

Smits, M., Ludden, G., Peters, R., Bredie, S. J., Van Goor, H., & Verbeek, P. P. (2022). Values that matter: a new method to design and assess moral mediation of technology. *Design issues*, *38*(1), 39-54. https://doi.org/10.1162/desi_a_00669

Steen, M. (2015). Upon opening the black box and finding it full: Exploring the ethics in design practices. *Science, Technology, & Human Values*, *40*(3), 389-420. https://doi.org/10.1177/0162243914547645

Sterling, B. (2005). Shaping things. MIT Press.

Swierstra, T., Stemerding, D., & Boenink, M. (2009). Exploring techno-moral change: the case of the obesity pill. *Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technology Developments.*, 119-138. https://doi.org/10.1007/978-90-481-2229-5_9

Toombs, A., Gross, S., Bardzell, S., & Bardzell, J. (2016). From Empathy to Care: A Feminist Care Ethics Perspective on Long-Term Researcher–Participant Relations. Interacting with Computers. https://doi.org/10.1093/iwc/iww010

Tromp, N., & Hekkert, P. (2018). *Designing for society: Products and services for a better world*. Bloomsbury Publishing.

Van Boeijen, A., Daalhuizen, J., & Zijlstra, J. (2020). *Delft design guide: Perspectives, models, approaches, methods*. BIS Publishers.

Van de Poel, I. (2011). Nuclear Energy as a Social Experiment. *Ethics, Policy & Environment 14*(3), 285-90. https://doi.org/10.1080/21550085.2011.605855

Van de Poel, I. (2013). Why new technologies should be conceived as social experiments. *Ethics, Policy & Environment, 16*(3), 352-355. https://doi.org/10.1080/21550085.2013.844575

Van de Poel, I. (2021). Design for value change. *Ethics and Information Technology*, *23*(1), 27-31. https://doi.org/10.1007/s10676-018-9461-9

Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago press.

Verbeek, P. P., & Tijink, D. (2020). *Guidance Ethics Approach: An Ethical Dialogue About Technology with Perspective on Actions. Den Haag, The Netherlands: Platform for the Information Society. https://research.utwente.nl/en/publications/guidance-ethics-approach-an-ethical-dialogue-about-technology-wit*

Vermaas, P. E., Hekkert, P., Manders-Huits, N., & Tromp, N. (2015). Design methods in design for values. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.) *Handbook of Ethics, Values and Technological Design* (pp. 179-202). Springer.

Van der Weij, F., Steinert, S., Van de Poel, I., Alleblas, J., Melnyk, A., & De Wildt, T. (2023). Value Change and Technological Design. *IEEE Technology and Society Magazine*, *42*(3), 25-32. https://ieeexplore.ieee.org/document/10260697

Whitbeck, C. (1996). Ethics as design: Doing justice to moral problems. *Hastings Center Report*, *26*(3), 9-16. https://doi.org/10.2307/3527925

Winner, L. (1980). Do artifacts have politics?. *Daedalus, 109*(1): 121-136. http://www.jstor.org/stable/20024652

About the Authors:

Deger Ozkaramanli is Assistant Professor in Human-Centered Design at Delft University of Technology. Her research is on design methods and ethics. She is the founding convener of DRS' SIG on Design Ethics.

Merlijn Smits is Associate Professor Human Centered Design at Saxion University of Applied Sciences. In her research, she combines design with ethics and philosophy of technology to design for values and user well-being, particularly in healthcare.

Maaike Harbers is Professor of Applied Sciences in AI & Society at Rotterdam University of Applied Sciences. She researches the ethical and societal aspects of AI, with a focus on the role of designers.

Gabriele Ferri is Assistant Professor at the Eindhoven University of Technology, focusing on Design Justice and Transforming Practices. His work leverages community-led design to address design-driven injustices and foster equitable alternatives. He engages in community collaborations and mentors students in justice-oriented projects.

Michael Nagenborg is an Associate Professor for Philosophy of Technology at the University of Twente. He is the co-director of the Philosophy of the City Research Group and a convener of DRS SIG on Design Ethics.

Ibo van de Poel is Professor in Ethics and Technology at Delft University of Technology. His expertise include value change, design for values, ethics of disruptive technologies, and moral responsibility. He has currently an ERC Advanced grant on Design for Changing Values.