Design for values

Van de Poel, Ibo

DOI
10.4337/9781035310685.00043

Publication date
2024

Document Version
Final published version

Published in
Handbook of Technology Assessment

30. Design for values

Ibo van de Poel

INTRODUCTION

Technology is based on, and affects, human values. For example, the use of various energy technologies affects the realization of such values as sustainability, human well-being, safety and (intergenerational) justice. Conversely, the design of technologies is always at least implicitly based on value-laden considerations. These may be practical, financial or economic considerations, but also morally motivated considerations and values.

Design for values aims at systematically accounting for values during the design of new technological products, services and systems. Here, I use ‘design for values’ as an umbrella term for a number of approaches that aim at integrating values of moral importance in the design of new technological products, services and systems in a systematic and proactive way. This includes approaches such as value sensitive design (Friedman/Hendry 2019), value-based engineering (Spiekermann 2015) and ethics-by-design, as well as approaches that are aimed at more specific values, such as privacy-by-design. These approaches are characterized by three features, namely: (1) an emphasis on values of moral importance; (2) the aim to proactively address values in the design of new technological products, services and systems; and (3) systematic attention to values throughout the design process.

In the literature on value sensitive design, values have been defined as ‘what is important to people in their lives, with a focus on ethics and morality’ (Friedman/Hendry 2019, p. 24). As a general and broad characterization, this will do for the current purpose. One thing to be aware of when using this definition is that there might be a difference between what people consider important in life, and what is morally valuable. For example, some people may find it important to treat people of a different gender, race or age differently than themselves, but that does not make discrimination morally valuable. Rather, values seem to denote what we have reason to value, or should value from a normative or moral point of view, but what people may not actually always value. Such a normative of morally motivated notion of value seems to motivate design for values rather than what people just consider important in life.

EMERGENCE, HISTORY AND THEORETICAL BACKGROUND

The value sensitive design (VSD) approach has been developed since the 1980s by Batya Friedman and colleagues in the field of human–computer interaction. Since then, the approach has also been applied to other technologies, such as energy technologies. Simultaneously, a number of somewhat similar approaches have developed in other fields of study, such as value-based engineering (Spiekermann 2015). More recently, we have seen a surge of ethics-by-design approaches, in response to ethical issues raised by, for example, specific technologies such as artificial intelligence, as well as X-by-design approaches, where ‘X’ is a specific value, such as privacy. In this chapter, these approaches are treated under the umbrella of 'design for values' (van den Hoven et al. 2015).

Although most design for values approaches were initially developed outside the field of technology assessment (TA), they have similar goals and there are some clear communalities, such as an emphasis on the social consequences of technology and the aim to democratize (see chap. 26) and to improve technological development (see chap. 42). Design for values may therefore now be seen as part of the repertoire of TA approaches that aim at constructively influencing technological development for the good, which also include such approaches as constructive TA (Schot/Rip 1997; see chap. 27) and real-time TA (Guston/Sarewitz 2002). What sets these approaches apart from traditional TA is the shift from assessment to constructively steering technological development, and attention by more actors than just governmental or societal actors (see chapters 7 and 36). Specific to design for values is the emphasis on the design of specific products, systems and services (rather than general technologies), and the emphasis on moral values (see chap. 25).

An important theoretical assumption behind design for values approaches is that technology is not value-neutral, but value-laden (Winner 1980). This means that the social and ethical ramifications of technologies depend not only on how they are used, but also on how they have been designed. In other words, it is possible to embed positive moral values through design in technology. Such embedded values are more likely to be realized during use, for example because they are the values that will be realized if a technology is used according to its original use plan (van de Poel/Kroes 2014), or because the technology affords actions that help to realize these values (Klenk 2021).

THE DESIGN FOR VALUES PROCESS

This section discusses the main elements of a design for values process, by paying attention to different activities or phases that might be distinguished , that is: (1) stakeholder analysis and value elicitation; (2) conceptualizing and specifying values; (3) identifying and dealing with value conflicts; (4) prototyping and testing; and (5) ongoing monitoring and value change.

Stakeholder Analysis and Value Elicitation

The first step in designing for values would typically be to elicit the relevant values for which a product, service or system needs to be designed. Sometimes, the focus might be on one specific value, for example in privacy-by-design approaches; but typically, a range of values are relevant for the design of technology.

The relevant values for which technology is to be designed can be established in a number of ways. One may look at relevant codes of conduct or legal rules and regulations that might stipulate relevant values. One might also investigate what the expected or potential impacts are of a technology, for example through a more traditional TA study, and so define what values should be taken into account in design.

Typically, design for values also requires stakeholder analysis and investigation into how stakeholders may be impacted upon and which values they deem relevant. In the literature on VSD, often a distinction is made between direct and indirect stakeholders (Friedman/Hendry 2019). Direct stakeholders are users and others that directly interact with the technology; indirect stakeholders may not directly interact with the technology but may somehow be impacted by it.

If design for values is to be aimed at realizing moral values, as suggested in the introduction, relevant values expressed by stakeholders cannot be taken at face value. Instead, the focus should be on for which values reasons exist to pursue them. This does not mean, however, that it is up to the designers, or their clients, to define what moral values should be pursued in a design for values process. Rather, it means to critically scrutinize whether stakeholders' values really correspond to moral values, and to be open to other moral values than those mentioned by stakeholders. This scrutinizing can take the form of a collective moral deliberation process (van de Poel/Zwart 2010).

Conceptualizing and Specifying Values

Values are too general and abstract to guide the design process directly. Therefore, they need to be conceptualized and specified. Conceptualization of values is 'the providing of a definition, analysis or description of a value that clarifies its meaning and often its applicability' (van de Poel 2013, p. 261). Different conceptualizations of a value may result in different conceptions of it. Specification of values refers to the translation of values into more specific norms and design requirements that can guide the design process.

Whereas conceptualization is largely independent from the specific context in which a value is applied, specification is context-specific, and aims at specifying what the value means in the specific context in which it is applied. As an example of conceptualizing values, consider the value of 'privacy', and in particular 'informational privacy', that is, privacy relating to information (data) about people. In the moral and legal literature, one can find at least three different conceptions of privacy.

A first conception of privacy understands it largely in terms of secrecy (see Warnier et al. 2015): a person's informational privacy is guaranteed if information about that person is secret, that is, it is not known to others. A second conception understands privacy in terms of control, that is, privacy requires that people have control over what information they share with whom, and for what purposes (e.g., Menges 2021). This is often expressed in terms of 'informed consent', for example in the European regulation on privacy, the GDPR (General Data Protection Regulation). Informed consent in this context means that people can freely decide what data to share, or not share, after they have been (fully) informed about how the data they share will be used and for what purposes. A third conception understands privacy in terms of contextual integrity (Nissenbaum 2004). This refers to the idea that people would typically find it appropriate to share certain data with some people, but not with others. For example, most people find it fitting to share medical data with their doctor, but not with their employer. Data sharing is thus to be guided by certain contextual norms that guarantee contextual integrity.

These three conceptions of privacy are likely to lead to different design decisions if privacy is considered a leading value in design. Consider, for example, the design of an e-health app. If privacy is understood as secrecy, such an app should collect as little (personal) data as possible (to fulfil its main functions). If possible, data is to be stored anonymously and to be deleted after some time. If privacy is understood in terms of control and informed consent, it becomes important to allow users to decide what data to share with whom. For example, a database structure is required to keep track of such decisions, as well as an interface that allows users to make such decisions, and to revise them if they wish at a later time. If privacy is understood as contextual integrity, it is mandatory to design the app in such a way that data streams can be regulated, and certain contextual norms about data sharing need to be established beforehand and to be built into the design of the app.

As can be seen from this example, different conceptions of a value are likely to lead to somewhat different design requirements. This brings us to the process of specification. Specification will usually assume a certain conception of a value, and will then translate that into more specific design requirements. Van de Poel (2013) proposes the values hierarchy as a tool for specifying values in a design context. A values hierarchy consist of three basic levels: values, norms and design requirements. Each level may be subdivided in two or more sublevels.

Values are here seen as the most general description of what is (morally) valuable and desirable. These can then be translated into more specific contextual norms that articulate what it means to strive for, or respect, a certain value in a given context. For example, aesthetics (or beauty) is a value in architecture. It may be translated into general norms such as that a building should be aesthetically pleasing or that it should follow the norms of a certain architectural tradition (such as classicism); the latter is clearly based on a more specific conception of aesthetics. These norms may subsequently be further translated into design requirements.

A values hierarchy may be construed top-down or bottom-up. In the first case, it follows the process described above. In the second case, it starts with design requirements and poses the question: For the sake of what are these striven for? In this way, underlying norms and values of a design may be reconstructed, even if they have not initially been explicitly articulated. In this way, the values hierarchy allows for articulating values and design requirements in an iterative fashion.

Identifying and Dealing with Value Conflicts

Although design for values may sometimes be focused on one value, it would typically involve a multiplicity of values. Even if, for example, a safe-by-design approach is followed, it would usually be irresponsible to ignore other relevant values, such as sustainability or justice, in the design process.

The multiplicity of values may result in a number of value conflicts surfacing in the design process. These value conflicts may take a number of shapes that are important to distinguish. One important distinction is that between cases in which two values are conflicting, and cases in which stakeholders have conflicting views on certain values (van de Poel 2021b).

As an example of the first, consider the design of a car that needs to be safe and sustainable (van Gorp 2005). One way in which cars can be designed in a more sustainable way is by reducing their weight so that they consume less energy (for the same distance driven). However, lowering the weight of the car will make it more difficult to incorporate some safety features, such as airbags or ABS (anti-lock braking system). So here the designers are facing a conflict between two values.

This case is to be distinguished from cases in which stakeholders disagree about values. In the car design case, stakeholders may disagree about whether sustainability is a relevant value, and if so, how important it is (compared to safety). More generally, stakeholder conflicts about value may surface in at least four different ways, namely as conflicts about: (1) the relevance of a value for a certain design task; (2) the relative importance of values; (3) the best conception of a value; and (4) the specification of a value for a given design task. These distinctions are relevant because they may require somewhat different strategies for addressing the value conflict.

In the literature, various strategies for dealing with conflicting values in design have been proposed (van de Poel 2015). When it comes to dealing with conflicts between values (rather than between stakeholders), four main types of strategies may be distinguished. The first type is maximizing or optimizing approaches that try to find the best design by making values somehow commensurable. Examples are cost–benefit analysis and multiple criteria analysis. A second possibility is satisficing, which means that a design is sought that is satisfactory, for example by defining threshold requirements that should be minimally met for each relevant value. A third approach is respecification. This approach employs the fact that a value can often be specified in a number of different ways. By trying out other adequate value specifications of the relevant values, it might be possible to come to a set of design requirements that is non-conflicting and still respects the underlying values. A fourth approach might be called innovation: it looks at finding new, not yet existing designs that avoid the initial value conflict.

When it comes to value conflicts between stakeholders, somewhat different strategies may be followed (van de Poel 2021b). One strategy is similar to what was called innovation: it tries to find a design that is optimal, or at least acceptable to all stakeholders. This may, for example, be done by using the method of value dams and value flows from VSD: a value flow is a design feature that is considered desirable by most, if not all, stakeholders and it should usually be included in the design; a design dam is a feature that is (strongly) opposed by (some) stakeholders and should therefore be avoided (Miller et al. 2007).

Value conflicts between stakeholders may also be resolved through some form of negotiation and compromise. There are, however, pitfalls to such an approach. One issue is that not all stakeholders may be equally visible and powerful; a negotiated compromise may thus reflect the views of powerful stakeholders, while neglecting those of marginalized stakeholders or future generations. Negotiation may thus lead to designs that are accepted by the most powerful stakeholders, but are nevertheless morally unacceptable.

Value conflict may also be the subject of deliberation and reflection among stakeholders. Through such deliberation one may try to come to what the political philosopher John Rawls (2001) has called an overlapping consensus. An overlapping consensus requires agreement on some issues, but not full agreement about underlying values or moral theories. In the case of design, stakeholders may come to agree about a set of design requirements for a technology even if they disagree – to some extent – about the relevant values, their relative importance or how to conceptualize these values.

Prototyping and Testing

Just like traditional design, design for values requires prototyping and testing of design proposals. With respect to values, prototyping and testing is important to validate whether a proposed design solution indeed embeds the intended values. One way to do this might be to test out whether, when a prototype is used properly, that is, according to its use plan, it indeed realizes its intended value (van de Poel/Kroes 2014). Of course, one might also want to test how robust a design is in realizing or respecting the intended values when used in unintended or unexpected ways. Prototyping and testing might also be used to see whether a design solution perhaps has unintended consequences, which may require taking additional values into account and revising the currently proposed design solution.

Ongoing Monitoring and Value Change

Often, the social consequences of – particularly new – technologies only become fully clear when they are already in use. That is to say, despite best efforts to anticipate the social consequences of new technology, and to proactively design for values, new technologies may have unexpected consequences that arise during the use phase. Therefore, ongoing monitoring of designed products, services and systems is required.

For example, products or services designed for sustainability may not only consume less energy, but also for that reason may be cheaper to use, which may – unintentionally – result in increased use, so cancelling out the positive sustainability effect (the so-called rebound effect). To address such unintended consequences (see chap. 5), a redesign might be required, or complementary measures might need to be taken.

One possibility that should explicitly be considered is that relevant values change over time. Van de Poel (2021a) distinguishes between five types of value change that may be relevant when designing for values. First, new values may emerge, or become more prominent, in society that are also relevant for the design of a particular technology. Think of the increasing societal emphasis on sustainability in the second half of the twentieth century. Second, a value that already existed but was deemed irrelevant for the design of a particular technology might become relevant. This may particularly happen if a technology has unexpected consequences. A third possibility is that the relative importance of values changes over time. A fourth possibility is that the (best) conceptualization of a value changes over time. For example, Google Glass required attention for spatial privacy, in addition to informational privacy, as it turned out that people perceived the device as a privacy intrusion when it was not used to record (Kudina/Verbeek 2019). Fifth, the specification of a value might change. This may, for example, be due to technological developments or new technological possibilities. For example, the specification of safety for vessels made from plastics and composites is quite different from that for steel vessels (van Gorp 2005).

METHODS AND APPROACHES

A host of methods is now available for design for values. Many of these are usual methods or tools from social science or design, although there are also a few tools and methods that have been developed specifically for design for values approaches.

Friedman and Hendry (2019) list 17 methods available to VSD: (1) stakeholder analysis; (2) stakeholder tokens; (3) value source analysis; (4) co-evolved technology and social structure; (5) value scenario; (6) value sketch; (7) value-oriented semi-structured interview; (8) scalable assessment of information dimensions; (9) value-oriented coding manual; (10) value-oriented mock-up, prototype or field deployment; (11) ethnographically informed inquiry on values and technology; (12) model for informed consent online; (13) value dams and flows; (14) value sensitive action-reflection model; (15) multi-lifespan timeline; (16) multi-lifespan co-design; and (17) envisioning cards.

This list is not exhaustive, however. When it comes to value elicitation, one might also think of a method such as Q-methodology that helps to find different stakeholder perspectives on an issue (O'Leary et al. 2013). This methodology helps to find the relevant stakeholder perspectives that should be taken into account in design for values, and may therefore have added value compared to more traditional stakeholder analysis methods. Also, text-mining tools might be of interest in value elicitation, or to investigate how values change over time (de Wildt et al. 2022; see also chap. 39).

Methods and approaches from the field of design are also useful for design for values. One might think of design approaches that help to structure the design process in different phases, but also of more specific approaches such as participatory design (e.g., Schuler/Namioka 1993). Participatory design aims at involving a broader range of stakeholders in the actual design process, not just in defining the design problem.

APPLICATIONS

Friedman and Hendry (2019, Chapter 4) discuss 11 applications of VSD; van den Hoven et al. (2015) discuss applications of design for values for different values and different application domains. Other insightful applications include civil healthcare drones (Cawthorne 2024); biorefineries (Palmeros Parada 2020); digital government collaborative platforms for sustainability (Sapraz 2023); and care robots (Poulsen 2022).

LIMITATIONS AND OBJECTIONS

A good overview of some criticisms and limitations of particularly VSD can be found in Davis and Nathan (2015). Here, three main criticisms and related limitations are discussed.

A first criticism is that not design but use determines what values are being realized through technology (see Albrechtslund 2007). There is certainly some truth in this criticism, in the sense that design choices do not fully determine use and the realization of values. Still, design choices matter: they determine what action possibilities are available to users, and enable and constrain the realization of certain values.

A second criticism is that design for values may result in a technological fix, in the sense that it claims to solve social and moral problems through technology. If technology really solves moral problems (and solves the right problems), there might be nothing wrong with a technological solution (van den Hoven et al. 2012). However, technological solutions often come with unintended consequences (see chap. 5), and too much focus on technology might lead to a focus on solvable problems rather than the 'real' problems.

A third criticism has been that design for values lacks a normative basis (Manders-Huits 2011). Interestingly, others have voiced what seems to be the opposite criticism, namely that it assumes universal values too much (Borning/Muller 2012). These opposing criticisms may reflect a broader issue, namely that design for values needs to operate in a context of value disagreement and moral uncertainty. In which case, it would seem undesirable that design for values as an approach commits to one specific moral theory.

CHALLENGES AND FUTURE DEVELOPMENTS

There has been considerable progress in developing design for values approaches, particularly through the work in VSD, during the last three decades. There are still, however, many challenges. One challenge is that there do not seem to be many good examples in which design for values approaches have been applied in practice, addressing a multiplicity of values from the start to the end of a design process, and where their application has been systematically evaluated. A second challenge is that the emphasis has often mainly been on the early design phases, such as value elicitation, value specification and dealing with value conflicts, and less on the complexities and pitfalls of translating values into technological choices and realizing values in practice. More attention is needed to the later phases of the design process, and practices of using, operating and maintaining technology with an eye to values. A final challenge relates to validation. There are currently no good methods for validating whether a design indeed embeds the intended values. This is particularly important because of the recent upsurge in approaches such as privacy-by-design, safe-by-design and ethics-by-design. Without proper validation methods, it is very hard to judge whether these methods fulfil their promises, but one thing that can already be said is that it will never be possible to resolve all privacy, safety or ethical issues through design. This is of course not a reason not to design for such values; but it might be better to speak of design for privacy (safety, ethics) rather than privacy-by-design.

ACKNOWLEDGEMENT

This chapter is part of the ValueChange project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, under grant agreement No. 788321.

REFERENCES

Albrechtslund, A. (2007): Ethics and technology design. Ethics and Information Technology 9: 63–72.
Borning, A.; Muller, M. (2012): Next steps for value sensitive design. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Austin, TX, ACM.
Cawthorne, D. (2024): The Ethics of Drone Design: How Value-Sensitive Design can Create Better Technologies. New York, Routledge.
Davis, J.; Nathan, L.P. (2015): Value sensitive design: applications, adaptations and critiques. In: van den Hoven, J.; Vermaas, P.; van de Poel, I. (eds): Handbook of Ethics and Values in Technological Design. Dordrecht, Springer: 11–40.
de Wildt, T.E.; van de Poel, I.R.; Chappin, E.J.L. (2022): Tracing long-term value change in (energy) technologies: opportunities of probabilistic topic models using large data sets. Science, Technology, and Human Values 47: 429–458.
Friedman, B.; Hendry, D. (2019): Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge, MA, MIT Press.
Guston, D.H.; Sarewitz, D. (2002): Real-time technology assessment. Technology in Society 24: 93–109.
Klenk, M. (2021): How do technological artefacts embody moral values? Philosophy and Technology 34: 525–544.
Kudina, O.; Verbeek, P.-P. (2019): Ethics from within: Google Glass, the Collingridge Dilemma, and the mediated value of privacy. Science, Technology, and Human Values 44: 291–314.

Manders-Huits, N. (2011): What values in design? The challenge of incorporating moral values into design. Science and Engineering Ethics 17: 271–287.
Menges, L. (2021): A defense of privacy as control. Journal of Ethics 25: 385–402.
Miller, J.K.; Friedman, B.; Jancke, G. (2007): Value tensions in design: the value sensitive design, development, and appropriation of a corporation's groupware system. Proceedings of the 2007 International ACM Conference on Supporting Group Work. Sanibel Island, FL, ACM.
Nissenbaum, H. (2004): Privacy as contextual integrity. Washington Law Review 79: 119–157.
O'Leary, K.; Wobbrock, J.O.; Riskin, E.A. (2013): Q-methodology as a research and design tool for HCI. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Paris, Association for Computing Machinery.
Palmeros Parada, M.D.M. (2020): Biorefinery Design in Context: Integrating Stakeholder Considerations in the Design of Biorefineries. PhD thesis, TU Delft.
Poulsen, A. (2022): The Investigation of A New Care Robot Design Approach for Alleviating LGBT+ Elderly Loneliness. PhD Thesis, Bathurst (AUS), Charles Sturt University.
Rawls, J. (2001): Justice as Fairness. A Restatement. Cambridge, MA, Belknap Press of Harvard University Press.
Sapraz, M. (2023): An E-Government Design Science Research in Sri Lanka: Facilitating Collaboration between the Government and Citizens for Environmental Sustainability. PhD thesis, Stockholm University, Department of Computer and Systems Sciences.
Schot, J.; Rip, A. (1997): The past and future of constructive technology assessment. Technological Forecasting and Social Change 54: 251–268.
Schuler, D.; Namioka, A. (eds) (1993): Participatory Design: Principles and Practices, Hillsdale, NJ, Lawrence Erlbaum.
Spiekermann, S. (2015): Ethical IT Innovation: A Value-Based System Design Approach. Boca Raton, FL, CRC Press.
van de Poel, I. (2013): Translating values into design requirements. In: Mitchfelder, D.; McCarty, N.; Goldberg, D.E. (eds): Philosophy and Engineering: Reflections on Practice, Principles and Process. Dordrecht, Springer: 253–266.
van de Poel, I. (2015): Conflicting values in design for values. In: van den Hoven, J.; Vermaas, P.E.; van de Poel, I. (eds): Handbook of Ethics, Values, and Technological Design. Dordrecht, Springer Netherlands: 89–116.
van de Poel, I. (2021a): Design for value change. Ethics and Information Technology 23: 27–31.
van de Poel, I. (2021b): Values and design. In: Michelfelder, D.P.; Doorn, N. (eds): Routledge Handbook to Philosophy of Engineering. Abingdon, Routledge: 300–314.
van de Poel, I.; Kroes, P. (2014): Can technology embody values? In: Kroes, P.; Verbeek, P.-P. (eds): The Moral Status of Technical Artefacts. Dordrecht, Springer: 103–124.
van de Poel, I.; Zwart, S.D. (2010): Reflective equilibrium in R&D networks. Science, Technology and Human Values 35: 174–199.
van den Hoven, J.; Lokhorst, G.-J.; van de Poel, I. (2012): Engineering and the problem of moral overload. Science and Engineering Ethics 18: 143–155.
van den Hoven, J.; Vermaas, P.E.; van de Poel, I. (eds) (2015): Handbook of Ethics and Values in Technological Design. Sources, Theory, Values and Application Domain. Dordrecht, Springer.
van Gorp, A. (2005): Ethical Issues in Engineering Design. Safety and Sustainability. Delft, Simon Stevin Series in the Philosophy of Technology.
Warnier, M.; Dechesne, F.; Brazier, F. (2015): Design for the value of privacy. In: van den Hoven, J.; Vermaas, P.E.; van De Poel, I. (eds): Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains. Dordrecht, Springer: 431–445.
Winner, L. (1980): Do artifacts have politics? Daedalus 109: 121–136.