Which Method for Engineering Concepts and Technologies?

Irene Olivero, Polytechnic University of Milan (Italy)
International Conference on Computer Ethics: Philosophical Enquiry (CEPE) 2023, Chicago, IL

Keywords: pragmatic method, socially disruptive technologies, conceptual engineering, deepfakes, function

Extended Abstract

Artifacts shape our life-world in several significant ways, and we rarely (if ever) find ourselves in a situation where no artifact is around. Surprisingly, despite their impact on our lives, they have not paid much philosophical attention. Recently, with the introduction and progressively increasing use of Artificial Intelligence, we find ourselves more and more surrounded and interacting with a particular kind of artifact, namely, artificial intelligence. AI is the common denominator of the so-called "socially disruptive technologies" (e.g., virtual assistants, deepfakes, self-driving vehicles, etc.). These kinds of artifacts deserve a lot of philosophical (and non) attention due primarily to the moral qualms and social challenges they bring about. Some of these socially disruptive technologies (SDTs) bring about so many issues that we need to consider whether it would be more beneficial to get rid of them or (at least) redesign some of their components. In this talk, I aim to call attention to such crucial entities by introducing a possible way to highlight and, at the same time, avoid the problems they bring about.

SDTs based on AI are increasingly present in our lives and transform them on social, ontological, and conceptual levels. Regarding the conceptual level, suffice to think about how artificial intelligence is reshaping our ways of thinking about home (due to virtual assistants), friendship (because of the introduction of robot companions), driving (from Google Maps to the Tesla), to mention a few. It seems crucial, then, that the recent "research field in analytic philosophy that focuses on how to assess and improve our representational devices" (Isaac 2021, 2053), i.e., conceptual engineering (CE), gets more involved with what happens in the field of technology. On the other hand, given their impact on our lives, we must be concerned about the SDTs' ethical, social, and ontological implications. I here argue that there may be a way for the two areas to help each other.

Given the "engineering" metaphor, it seems almost natural to think that the work in CE gets more focused on the role of technology in conceptual change and conceptual disruption. Despite the metaphor, instead, conceptual engineering has given little attention to these themes. Interestingly, by contrast, Amie L. Thomasson's approach (2020) to conceptual engineering takes the "engineering" metaphor seriously by drawing her method for assessing and revising concepts from the process adopted in the engineering and construction of artifacts and technologies. Thomasson advocates for adopting a pragmatic approach to CE that takes "the function of our (ranges of) concepts as playing a central role (2020, 440). There are three crucial passages in this method. We should: employ reverse engineering, i.e., look at the genealogy of the concept under examination (cf. Plunkett 2016) and try to determine what it does and can do; identify the function(s) (if any) the concept in question should serve and is to serve given the goals and purposes we have; finally, constructively engineer the concept at stake, given the function(s) we need it to perform (cf. also Haslanger 2000).

I hold that this pragmatic method for CE may be of inspiration for the philosophy of technology when it comes to "engineering" the ontology and metaphysics of the emerging SDTs, for it can also take into account the ethical and social concerns these technologies bring about. By applying Thomasson's method for CE not to concepts but to SDTs, we can attempt to investigate questions about these emerging technologies that highlight even more why they are so important to us and whether and how we should revise them. Consider, for example, deepfakes. We may ask about these: what is the function (if any) of deepfakes? What function(s)

(if any) do we want deepfakes to serve? How can we change the DLA, i.e., what should we change in the set of rules which constitute the basis of deepfakes so that they can perform the function they should perform? As with concepts, answering these questions may help us decide how to re-engineer (or whether to eliminate some of them, e.g., were they not to serve the function we want them to) deepfakes in particular, but also, mutatis mutandis, any of our emerging SDTs. This method may help us see what we should and should not keep about these technologies.

To give a concrete example, take the following parallelism. We legitimately keep the concept of marriage because of its purposes, but we reshape it to include same-sex relationships to serve those aims better. Similarly, on the one hand, we may want to keep the artificial intelligence that makes the creation of something like Microsoft's Rembrandt and Kennedy's speech (cf. Floridi 2018) possible. On the other hand, we may also want to re-engineer that technology to limit its scope and avoid using it for criminal or evil purposes (e.g., as it happens with deceptive, unauthorized deepfakes).
As with the CE method, the advantage here is that we would re-evaluate our SDTs a posteriori after we already know most of their possible uses and functions. With artifacts and emerging technologies, we have to deal with discoveries and "surprises" constantly: it often happens that a certain artifact that was invented for, or that we thought performed, a certain function, turns out also to serve well or even better, other functions (e.g., aspirins, cable-wires, vacuum tubes, etc.). However, it can also turn out that it can be used for evil purposes or has functions that have bad outcomes. Yet, what matters in all these cases is what to do, what we should do were these circumstances to occur. Among other methods (cf. Reijers et al. 2018), the pragmatic method proposed can address the problems that arise and could not be foreseen during their invention and production.

References

Baker L. R. (2007), The Metaphysics of Everyday Life: An Essay in Practical Realism, Cambridge University Press, Cambridge.

Bloom P. (1996), "Intention, History, and Artifact Concepts," Cognition 60, pp. 1-29.

Burge T. (2003). "Concepts, conceptions, reflective understanding: Reply to Peacocke." In Martin Hahn and B. Ramberg (eds.), Reflections and Replies: Essays on the Philosophy of Tyler Burge. MIT Press.

Carey S., Conceptual Change in Childhood, MIT Press, Cambridge 1985.

Carrara M. and D. Mingardo (2013), "Artifact Categorization. Trends and Problems," Review of Philosophy and Psychology 4, pp. 351-373.

Carrara M. and P. Vermaas (2009), "The fine-grained metaphysics of artifactual and biological functional kinds," Synthese 169, pp. 125-143.

Cummins R. (1975), "Functional analysis," The Journal of Philosophy 72, pp. 741-765.
Dipert R. R. (1993), Artifacts, Art Works, and Agency, Temple University Press, Philadelphia.

Elder C. (2007), "On the Place of Artifacts in Ontology," in E. Margolis and S. Laurence (eds.), Creations of the Mind: Essays on Artifacts and their Representation, Oxford University Press, New York, pp.

33-51.

Evnine S. (2016), Making Objects and Events, Oxford University Press, Oxford.

Fallis, D. (2020), "The Epistemic Threat of Deepfakes," Philosophy and Technology 34 (4), pp. 623-643.

Floridi L. (2018), "Artificial intelligence, deepfakes and a future of ectypes," Philosophy and Technology 31 (3), pp. 317-321.

Grandy R. E. (2007), "Artifacts: Parts and Principles," in E. Margolis and S. Laurence (eds.), Creations of the Mind: Essays on Artifacts and their Representation, Oxford University Press, New York, pp. 18-32.

Haslanger S. (2012), Resisting Reality: Social Construction and Social Critique. Oxford University Press.

Haslanger S. (2020), "How Not to Change the Subject." In Teresa Marques and Åsa Wikforss (eds.), Shifting Concepts, Oxford. University Press, pp. 235-259.

Hilpinen R. (1999), "Artifact," The Stanford Encyclopedia of Philosophy, Stanford University Press, Stanford. Hilpinen R. (1992), "On artifacts and works of art," Theoria 58, pp. 58-82.

Houkes W. and P. E. Vermaas (2010), Technical Functions: On the Use and Design of Artefacts, Springer, Dordrecht.

Isaac M. G. (2021), "Which Concept of Concept for Conceptual Engineering?," Erkenntnis: An International Journal of Scientific Philosophy, pp. 1–25.

Jansen L. (2013), "Artefact Kinds Need Not Be Kinds of Artefacts", in Christer Svennerlind, Jan Almäng and Rögnvaldu Ingthorsson (eds.), Johanssonian Investigations. Essays in Honour of Ingvar Johansson on His Seventieth Birthday, Ontos., Heusenstamm, pp. 317-337.

Keil F. C. (1989), Concepts, Kinds, and Cognitive Development, MIT Press, Cambridge.

Keil F. C., M. Greif, and R. S. Kerner (2007), "A world apart: How concepts of the constructed world are different in representation and in development," in E. Margolis and S. Laurence (eds.), Creations of the Mind: Essays on Artifacts and their Representation, Oxford University Press, Oxford, pp. 231-245.

Kitcher P. S. (1993), "Function and design," in Midwest studies in philosophy, XVIII, ed. P. A. French, T. E. Uehling, H.K. Wettstein, University of Minneapolis Press, Minneapolis, pp. 379-397.

Margolis E. and S. Laurence (eds.), Creations of the Mind: Theories of Artifacts and Their Representation, Oxford University Press, New York 2007.

Margolis E. (1998), "How to Acquire a Concept," Mind and Language 13, pp. 347-369.

Matan A. and S. Carey (2001), “Developmental Changes within the Core of Artifact Concepts,” Cognition 78, pp. 1-26.

McLaughlin P. (2001), What Functions Explain, Cambridge University Press, Cambridge.

Millikan R. G. (1999), “Wings, Spoons, Pills, and Quills: A Pluralist Theory of Function,” The Journal of Philosophy 99, pp. 191-206.

Millikan R. G. (1993), “In Defense of Proper Function,” in White Queen Psychology and Other Essays for Alice, MIT Press, pp. 13-29.

Millikan R. G. (1984), Language, Thought and Other Biological Categories: New Foundations for Realism, MIT Press, Cambridge.

Mumford S. (1998), Dispositions, Clarendon Press.

Neander K. (1991), “Functions as selected effects: The conceptual analyst’s defense,” Philosophical Science 58, pp. 168-184.

Neander K., “The teleological notion of ‘function’”, Australasian Journal of Philosophy 69 (1991b), pp. 454-468.

Neander K. (2002), “Types of traits: The importance of functional homologues,” in Functions: New Essays in the philosophy of psychology and biology, ed. A. Ariew, R. Cummins, M. Perlman, Oxford University Press, Oxford, pp. 390-415.

Öhman C. (2020), “Introducing the pervert’s dilemma: a contribution to the critique of Deepfake Pornography,” Ethics Inf Technol 22, pp. 133–140.

Olivero I. (2019), “Function is Not Enough: An Externalist Defeat for Artifactual and Social Kind Terms,” Grazer Philosophishe Studien 96 (1), pp. 105-129.

Plunkett D. (2016), “Conceptual History, Conceptual Ethics, and the Aims of Inquiry: A Framework for Thinking about the Relevance of the History/Genealogy of Concepts to Normative Inquiry,” Ergo: An Open Access Journal of Philosophy 3, pp. 27-64.

Preston B. (2009), “Philosophical Theories of Artifact Functions,” in A. Meijers (ed.), Philosophy of Technology and Engineering Science, Elsevier, Amsterdam, pp. 213-233.

Preston B. (1998), “Why is a Wing Like a Spoon; A Pluralist Theory of Function,” The Journal of Philosophy 95, pp. 215-254.

Reijers W.; Wright D.; Brey P.; Weber K.; Rodrigues R.; O’Sullivan D. and Gordijn B. (2018), “Methods for Practising Ethics in Research and Innovation: A Literature Review, Critical Analysis and

Recommendations", Science and Engineering Ethics 24 (5), pp. 1437-1481.

Rini R. (2019). "Social media disinformation and the security threat to democratic legitimacy", NATO Association of Canada: Disinformation and Digital Democracies in the 21st Century, pp. 10-14.

Rips L. J. (1989), "Similarity, typicality and categorization," in S. Vosniadou and A. Ortony (ed.), Similarity and Analogical Reasoning, Cambridge University Press, Cambridge, pp. 21-59.

Risse M. and Kerner C. (2021), "Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfake Technology in Digital Lifeworlds," Moral Philosophy and Politics 8 (1), pp. 81-108.

Scheele M. (2006), "Functions and use of technical artefacts: Social conditions of function ascription," Studies in History and Philosophy of Science 37, pp. 23-36.

Scheele M. (2005), The Proper Use of Artefacts: A Philosophical Theory of the Social Constitution of Artefact Functions, Simon Stevin Series in the Philosophy of Technology.

Sperber D. (2007), "Seedless Grapes: Nature and Culture," in E. Margolis and S. Laurence (eds.), Creations of the Mind: Theories of Artifacts and Their Representation, Oxford University Press, New York, pp. 124- 137.

Thomasson A. L. (2020), "A Pragmatic Method for Conceptual Ethics," in H. Capellen, D. Plunkett, and A. Burgess (eds.), Conceptual Engineering and Conceptual Ethics, Oxford University Press, New York, pp. 435-458.

Thomasson A. L. (2014), "Public Artifacts, Intentions, and Norms," in M. Franssen et al., Artefact Kinds. Ontology and the Human-Made World, Springer, Dordrecht, pp. 45-62.

Thomasson A. L. (2007), "Artifacts and Human Concepts," in E. Margolis and S. Laurence (eds.), Creations of the Mind: Theories of Artifacts and Their Representation, Oxford University Press, New York, pp. 52- 73.

Thomasson A. L. (2003), "Realism and Human Kinds," Philosophy and Phenomenological Research 67, pp. 580-609.

Vega Encabo J. and D. Lawler (2014), "Creating Artifactual Kinds," in M. Franssen et al. (eds.), Artefact Kinds. Ontology and The Human-Made World, Springer, Dordrecht, pp. 105-124.

Waelen R. and Brey P. (2022), "Ethical Dimensions of Facial Recognition and Video Analytics in Surveillance", in Michael Boylan & Wanda Teays (eds.), Ethics in the AI, Technology, and Information Age, pp. 15.

Wiggins D. (2001), Sameness and Substance Renewed, Cambridge University Press, Cambridge.