

Computing Professionals for Social Responsibility: The Past, Present and Future Values of Participatory Design • Contributions

Designing Across Distributed Agency: Values, participatory design and building socially responsible AI

Stephen C Slota

Published on: May 25, 2020

Updated on: May 06, 2020

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Designing Across Distributed Agency: Values, participatory design and building socially responsible AI

By *Stephen C. Slota*.

School of Information, University of Texas at Austin.

Contact: steveslota@gmail.com

While data science is vociferous about its value, it is oddly silent about its values. Algorithmically-driven analysis of very large data sets, often falling under broader categories of machine learning (ML) and artificial intelligence (AI) has already become an integral part of our social fabric. We continually monitor ourselves, each other, and the environment in an increasingly data-intensive manner, and policy decisions are more and more often being made on the basis of the results of opaque, occulted (Slota, Slaughter and Bowker, 2020) or otherwise black-boxed analysis of very large, heterogeneous data sets. Our relationship with data is changing, both as individuals and as a society. Self-tracking and quantification provides a data-oriented mode of understanding our own behaviors, bodies, and lives. (Berson, 2015) As a society, we are rendering ourselves increasingly computable, with greater portions of our lives tracked, quantified, and analyzed as part of our daily activities. (Cheung, et al. 2017) Similarly, we are changing the nature of our society - creating new kinds of 'social facts' built around the ubiquity and presence of tracking, data collection, and the presentation of results of that collection to us. (Boullier, 2015) We are witnessing fundamentally new modes of social organization such as the widespread use of predictive analytics in a variety of fields, which can work to create new ways in which our behaviors and practices might be governed. (Foucault, 1977; Mackenzie, 2013; Lövbrand et al., 2012; Williamson, 2016; Dow-Schull, 2012) Even disregarding these external or societal orientations towards data and analytics, there is reason to believe that our own relationships, attitudes, and priorities around data collected about us is changing. For example, both the scope of the available technology and our social understanding of the nature of a right to privacy are fundamentally different from the situation fifty, or even twenty, years ago (Solove, 2015; Martin, 2014; Nissenbaum, 2009).

Emergent properties of large, connected heterogeneous data sets and the algorithms needed to effectively mine useful information from them are substantively different from data analytic techniques that work with intentionally-collected data and relatively confined research spaces. Data science is unique in its deviation from the traditional 'data lifecycle' that begins with observation and moves through storage, cleaning, analysis and revision (Borgman, 2007; Borgman et al., 2008). Rather than working from initial observations of the world, novel data science leverages trace data (Geiger and Ribes, 2011) data collected from a heterogeneous sources towards a variety of different uses, and data collected about individuals largely without their knowledge producing its outputs. As such the

variety and quantity of data as well as the often-counterintuitive outputs of algorithmically driven analysis such as machine learning make it significantly more difficult to predict harms; they enroll a substantively larger population of individuals being observed through increasingly varied avenues of data collection; and their conclusions often have the weight of knowledge without significant accounts of their uncertainty. Beyond all of this are issues arising from attempts to understand and assert the accountability of increasingly 'black-boxed' analytic techniques – particularly when the outcomes of research conducted using those techniques are used to inform policy, direct resources, and respond to emergency situations (Lehr and Ohm, 2017).

Data collection, description, and curation bears consequence as much as its analysis and implantation in the world. An expert system, say, that informs sentencing decisions, or the regime of sensing, analysis, and policy that would enable the broad development of autonomous vehicles, is better considered as a complex network rather than a single technology. As such, individual designers, as nodes in that network, would naturally have some difficulty in considering the broader impact of such design decisions, rendering a consequentialist approach to ethically sound system design intractable to all but relatively distant hindsight. Instead, the work of ethics across these systems is located across the negotiation of specific values, repeatedly, and across time. There is not a single decision or design point at which a system might be rendered transparent, but rather numerous moments at which the technology and its data, algorithmic, and policy infrastructure might be more inflecting towards transparent design. *However*, current legal and policy structures generally assume individual responsibility in a way that becomes problematic when agency and decision-making is distributed between individuals and AI/ML systems that inform or their decisions or act autonomously. From laws that assume a vehicle has a single driver in control to existing standards and regulation for the education and training of medical professionals, law and policy broadly operates on the assumption of the agency of the individual. AI and ML, in overt and occluded ways, has the potential to distribute or relocate that agency – and may already be doing so (Heer, 2019; Skeem and Lowencamp, 2016; Harcourt, 2008).

In its capacity to democratize the innovation process (Björgvinsson, et al. 2010) and bridge the gaps in design expertise across collaborative teams (Scariot, Heemann, and Padovani, 2012) participatory design has the potential to acknowledge and account for the relocation of agency in large scale data analytics in order to more enable a broader conversation about the values, operation, and limitations to the application of its outcomes. Socially responsible design of large-scale data analysis must build towards an infrastructure that supports transparent understanding of the distributed ecology of design, curation, and implementation that characterizes those systems that augment, automate, and inform human decision-making and work. Participatory design, with its roots in emancipation and activism (Schuler & Namioka, 1993) has substantial potential to address these 'knotted' issues (Jackson, et al. 2014) by encouraging group reasoning and in 'closing the gaps' of expertise between those

working directly on AI and ML systems and those who make use of their outputs. In drawing together the distributed and heterogeneous points of values-inflection that scaffold the social impacts of AI and ML, participatory collaboration across the design – and more importantly, interpretation and implementation – processes and infrastructures represents not only a substantive opportunity to encourage responsible and accountable design, but also may be a necessity in understanding and accounting for its impacts.

References

- Berson, Josh. 2015. *Computable Bodies: Instrumented Life and the Human Somatic Niche*. London: Bloomsbury.
- Björgvinsson, Erling, Pelle Ehn, and Per-Anders Hillgren. 2010. "Participatory design and democratizing innovation." In *Proceedings of the 11th Biennial participatory design conference*: 41-50.
- Borgman, C. L. 2007. *Scholarship in the digital age: Information, infrastructure, and the internet*. Cambridge, MA: The MIT Press.
- Borgman, C. L., Abelson, H., Dirks, L., Johnson, R., Koedinger, K. R., Linn, M. C., et al. 2008. *Fostering learning in the networked world: The cyberlearning opportunity and challenge, a 21st century agenda for the National Science Foundation*. Arlington: NSF Task Force on Cyberlearning.
- Boullier, Dominique. "The social sciences and the traces of big data: Society, opinion, or vibrations?." *Revue française de science politique (English Edition)* 65, no. 5-6 (2015): 71-93.
- Cheung, Cynthia, Matthew J. Bietz, Kevin Patrick, and Cinnamon S. Bloss. 2017 "Conceptualizations of Privacy Among Early Adopters of Emerging Health Technologies." In *Annals of Behavioral Medicine* 51: S2586.
- Dow Schüll, Natasha. 2012. *Addiction by design: Machine gambling in Las Vegas*. Princeton, NJ: Princeton University Press.
- Foucault, M. 1977. *Discipline and punish: The birth of the prison*. Vintage.
- Geiger, R. Stuart, and David Ribes. 2011. "Trace ethnography: Following coordination through documentary practices." In *System Sciences (HICSS)*, 2011: 1-10.
- Harcourt, B. 2007. *Against Prediction: Profiling, Policing and Punishing in an Actuarial Age*, Chicago: Chicago University Press.
- Heer, Jeffrey. 2019. "Agency plus automation: Designing artificial intelligence into interactive systems." *Proceedings of the National Academy of Sciences* 116, no. 6: 1844-1850.

Jackson, Steven J., Tarleton Gillespie, and Sandy Payette. 2014 "The policy knot: re-integrating policy, practice and design in cscw studies of social computing." In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*: 588-602.

Lehr, D., & Ohm, P. 2017. "Playing with the Data: What Legal Scholars Should Learn About Machine Learning." *UCDL Rev.*, 51: 653.

Lövbrand, E., Stripple, J., & Wiman, B. 2009. "Earth System governmentality: Reflections on science in the Anthropocene." *Global Environmental Change* 19 no. 1: 7-13.

Mackenzie, A. 2013. "Programming subjects in the regime of anticipation: Software studies and subjectivity." *Subjectivity* 6 no. 4: 391-405.

Martin, J. P. 2014. "Brief History of Privacy and Selected Electronic Surveillance Laws." *Cloud Computing and Electronic Discovery*. Martin, J. and Cendrowski, H. Eds. Hoboken, NJ: John Wiley and Sons: 37-54.

Nissenbaum, H. 2009. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Scariot, Cristiele A., Adriano Heemann, and Stephania Padovani. "Understanding the collaborative-participatory design." *Work* 41, no. Supplement 1 (2012): 2701-2705.

Schuler, Douglas, and Aki Namioka, eds. 1993. *Participatory design: Principles and practices*. CRC Press.

Skeem, Jennifer L., and Christopher T. Lowenkamp. 2016. "Risk, race, and recidivism: predictive bias and disparate impact." *Criminology* 54, no. 4: 680-712.

Slota, Stephen C., Slaughter, Aubrey and Bowker, Geoffrey C. "Chapter 1: The Hearth of Darkness: Living Within Occult Infrastructures." Lievrouw, Leah and Loader, Brian, eds. *Handbook of Digital Media and Communication*. Routledge, London. Forthcoming 2020.

Solove, D. J. 2015. "The meaning and value of privacy." In Roessler, B. and Mokrosinska, D. Eds. *Social Dimensions of Privacy: Interdisciplinary Perspectives* Cambridge, MA: Cambridge University Press: 71-82.

Williamson, Ben. 2016. "Digital education governance: data visualization, predictive analytics, and 'real-time' policy instruments." *Journal of Education Policy* 31, no. 2: 123-141.