Mitchell Sutika-Sipus

May 17, 2019

Submission to the UN Special Rapporteur, Philp Alston.

Review of Emergent Design Processes of Socio-Technical Interventions

Please publish on webite.

Summary

While much has been written on the role of digital technology in the deteriorated access to quality of information or the role of disinformation in democratic deterioration, and its impact on human rights, there has been little attention to the issue of ‘how technology is made.” Yet policy is not enough. Execution of values through a material can realize threat opportunity or disaster. Regardless of solution paths determined, a consideration must be given to ‘how’ we implement those socio-technical solutions. An overview is provided of emerging concerns and methods of design to counter these concerns is provided.

*Introduction*

Design matters more today than any time since WWII. With the transformation of computational technology and its saturation within daily life, there is a marked decline of empathy in social discourse, the polarization of beliefs, and the fragmentation of governance is tied to the creation and movement of information. The reproduction and distribution of information at light speed holds no relation to the creation of quality. Data is a poor substitute for nuance. The compression of a narrative into a number is no sin, and in some instances is even preferable, yet numbers alone are not truth. By consequence, the world is suffering from the demand for story tellers, truth speakers, and information leaders.

The pause for reflection in Design, and the pursuit of experiments on what design could be or how it is done, are essential at this time as our problems are computational, not industrial, and thus generate consequences in multiples. The creation of rich and interactive imagery has been democratized and the channels of engagement once restricted through commercial and purely social spaces are now blurred and globally distributed through YouTube and Instagram. Any individual can make anything, can become anyone, and possibly connect with anyone else – or one can use this same enablement toward destructive means.

Surprisingly, no art or design school in America is tackling the crisis of disinformation or the inequitable distribution of technology capabilities in the world. Though studied by RAND, the Pentagon, and US Department of State – no bureaucrat, politician, or social scientist has an answer to our digital malaise. There is much discussion of ethics and principles to inform machine learning, but market forces are not necessarily attuned to these discussions. Many ask technology companies to take responsibility – yet these companies can only rely upon their engineers who are otherwise embedded in an algorithmic arms race. However, interaction Design and user experience, as an expressive mediation between information creation and consumption, remains absent within the conversation on countering disinformation. The demand for design research to counter the obstructions of technology is great.

The Design Problems

The technologies that underlie current social problems in the creation and sharing of disinformation are generally common place forms of social media, messaging boards, email, and online advertising. The techniques, likewise, are not necessarily exceptional and are common practice within digital marketing and messaging campaigns by individuals and companies. However, three primary elements are listed:

#### Non-Transparency

The role of machine learning continues to expand within domains of social and political importance. Black box algorithms however continue to deny clarity on internal bias of decision making while bias within data continues to introduce problems that magnify within product adoption. Deep learning methods in particular are a challenge to assign accountability. (Hindman 2018). Transparency is particularly important when considering matters of personal data, the use of non-personal data to micro-target people, and the ability to reconfigure social environments through the manipulation of data tied to individual vulnerabilities (Lievens 2019).

#### Inaccessibility

Complimentary to the lack of transparency is the inequitable distribution of access to information creation. Large datasets about public and private life are concentrated among small elite groups (Bradshaw and Howard 2019). This problem is reinforced with black box algorithms and sophisticated computational architectures which distort the provenance of information, disarm individuals and spread information through bots, create endless echo chambers of messaging, and develop extrapolations of information that chaotically circulate through public spheres to manipulate public perceptions of the world.

#### Simulated Social Actors (Bots)

Social bots are computational simulations of human actors that can assist with computational tasks in substitute of a human actor. However, these bots can also exploit human vulnerabilities by manipulating shortcoming in human reasoning, by deploying cyber-attacks, or by generating fake online profiles that distort the public information landscape by creating and circulating false content. Bot use proliferates within election manipulation or high-stress political conditions, as the 2016 US election surfaced that 33% of all user accounts generating disinformation were in fact bots (Hindman 2018). In another example, Russian media uses negative news about the US to distract Russian citizens when its own economy is facing challenges (Field, et al. 2018).

Platform DesignSocial media platforms have democratized content creation via the design of decentralized systems with multiple contributors. While technological capacities exist to recognize disinformation more vigorous forms of self-governance are in demand (Hindman 2018). As it stands, a range of social computing design elements within social media platforms drives particular kinds of behaviors and empowers particular kinds of actors. In addition to the use of bots, a specialized form is called an *automated sociopathic actor*, a bot that is used to amplify information as it moves through network. Additionally, elements that act as *information recasting tools* craft and apply subtle variations to information and redistribute to drive the proliferation of subtle nuances, increase distrust, and drive animosity.

#### The Insular Demands of Machine Logic

Whereas all technologies may be used for to accomplish a given goal with positive or harmful intention, advanced artificial intelligence systems are different from other technologies as such technologies possess their own goals, objectives, in combination with advanced reasoning and extensibility (such as robotic hardware extensions), become technologies that can misuse themselves. In this manner, the logic of the machine is to optimize the world according to the demands of its goal, not human values. At this time, a highly advanced artificial general intelligence (AGI) does not exist, yet as our learning technologies become more sophisticated a viable threat begins to arise as an AGI system applies its energies to itself, recursively developing a superior intelligence. With extended intelligence and goal stability, such a machine becomes superior only to the needs of itself (Bostrom 2018).

*Emerging Approaches to Counter the Problems of Machine Logic and Machine Systems*

One approach gaining traction is the pursuit of human centered approaches to machine learning. Such an approach considers algorithms and interfaces in support of human goals, contexts, and ways of working to make machine learning tools and technologies more usable and possibly useful. Methods include permitting users to modify algorithms in real-time by working with the curation and supply of various types of training data, by changing their own human behavior, and permitting the opportunity for users to critique outputs (Gillies, et al. 2016).

A key element is for design methods to engage the co-adaptive processes of machine learning, in which a human is changing computer behavior, but the human also adapts to the tool or changes goals in relation to the tool (Gillies, et al. 2016). In this context, goal formation and reformation is a unique human ability in which human goals may exist arbitrarily, whereas computational goals are tied to criteria and the supportive architectures are not adaptive systems. For a computer to change its goals in the same manner as a human is an unlikely future (Bostrom 2018).

Another approach to materializing machine learning with the design process is a deeper reliance on metaphor and modelling. Within a co-design process, it is more effective for new terms and language to be created that references the goals and interests of the groups, rather than a reliance on technical terms. Likewise, this may dilute the necessity to rely on deep algorithmic expertise within the design process, and to permit designers and stakeholders to focus more closely on goal formation and problem definition. As clarity takes form, the design process can probe the interactions on machine learning requirements such as training, testing, and explore the impact of errors. It is furthermore advantageous for designers to adapt a data-driven mindset and to explore the implications of data creation, processing, and data relationships within the product, user experience, and consequences of interaction (Yang, Scuito, et al. 2018).

At this time there is little leadership among designers in understanding and working with machine learning. Education on the topic poor and the ability for students to penetrate the bounded reasoning of machine logic is a longstanding issue. Notably, the challenge is an old one. Early AI pioneers Marvin Minsky and Joseph Weizenbaum cited engagement between humans and machines as asymmetric, as human conversation is dependent upon social context, personal memories and experiences for humans to engage and understand one another within a range of similarity (Weizenbaum 1967). In contrast, a computational approach is essentially algebraic, assembling words according to sets of measurements, rules, and computations to generate language that is interpreted by the human. The machine offers no other form of meaning – for example, a machine does not “know” a problem, like a human, and interaction may be meaningful only as the human supplies the meaning through the layering of human thoughts including emotions and heuristics (Minsky 2007). For the human actor, the logic of the machine is alien, and to construct a technological user experience through machine learning demands new conceptions of interaction and methods of engagement.

*Emerging approaches to Design for Machine Learning*

At this time there is no clear research agenda regarding Machine Learning and design, as the technology is sufficiently new to that questions held by design are at the lowest levels of sophistication. User Experience research concerning ML has focused on the interaction of simulated agents, in particular the UX of conversations, turn taking, and effectiveness of agents within states of uncertainty. These agents have amplified human characteristics such as gender or ethnicity, and further maintain other conditional elements to reinforce human perception and beliefs. A possible future extension of research in this domain concerns UX, human behavior, and response to control autonomous vehicles or conduct additional real-world tasks via digital agent interactions (Dove, et al. 2017).

Machine learning also poses challenges to the design process. It is hard to envision and hard to conceive as material and thus hard to prototype. Reliance on sketches, plans, models and stories is evocative but the realization of learning and large-scale data driven implications is difficult. Many designers do not understand the limits of machine learning and thus the development of machine learning products is not a mature practice beyond a focus on the implementation of specific uses cases such as face recognition. While in recent years designers have advanced their abilities to work with software as material, the extended temporal processes of underlying machine learning has yet to take form within design, in part because the technology is dependent on data and tooling frequently unavailable (Dove, et al. 2017).

One approach gaining traction is the pursuit of human centered approaches to machine learning. Such an approach considers algorithms and interfaces in support of human goals, contexts, and ways of working to make machine learning tools and technologies more usable and possibly useful. Methods include permitting users to modify algorithms in real-time by working with the curation and supply of various types of training data, by changing their own human behavior, and permitting the opportunity for users to critique outputs (Gillies, et al. 2016). A key element is for design methods to engage the co-adaptive processes of machine learning, in which a human is changing computer behavior, but the human also adapts to the tool or changes goals in relation to the tool (Gillies, et al. 2016). In this context, goal formation and reformation is a unique human ability in which human goals may exist arbitrarily, whereas computational goals are tied to criteria and the supportive architectures are not adaptive systems. For a computer to change its goals in the same manner as a human is an unlikely future (Bostrom 2018).

For Design to make advances in designing with machine learning, it is suggested that designers explore the interplay of system 1 and system 2 thinking. The power of systems 2 reasoning can be offsetting for humans, whereas design opportunities exist to better communicate such power by helping humans anticipate errors, identified false positives, or communicate the limited perception of the problem within the computation. There is also the opportunity for the integration of human domain knowledge within the tool (Koch 2017). It is advised that designers search for alternative ways to represent machine learning dependencies on data in early prototypes and to search out new applications of machine learning. Lastly, designers must be educated on the ethics and social implications of machine learning technologies in the world (Dove, et al. 2017).

# Bibliography

Bostrom, Nick. 2018. "Interview with Nick Bostrom." In *Architects of Intelligence: The truth about AI from the people building it*, by Martin Ford, 97-116. Birmingham: Packt Publishing Ltd.

Bradshaw, Samantha, and Philip N. Howard. 2019. *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation.* Computational Propaganda Research Project, Oxford: University of Oxford Internet Institute.

Dove, Graham, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. "Design Innovation: Challenges for Working with Machine Learning as a Design Material." *Proceedings of the SIGHCI Conference on Human Factors in Computing Systems.* (CHI'17).

Field, Anjalie, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia. Tsvetkov. 2018. "Faming and Agenda-setting in Russian News: a computational analysis of intricate political strategies." *arXiv:1808.09386v2 [cs.CL] 29 Oct 2018.*

Gillies, Marco, Frederic Bevilacqua, Alexis Heloir, Fabrizio Nunnari, Wendy Mackay, Saleema Amershi, Bongshin Lee, et al. 2016. "Human-Centered Machine Learning." *Chi 2016 Workshop Summary.* San Jose: CHI. 3558-3565.

Hindman, Matthew. 2018. *Disinformation, 'Fake News' and Influence Campaigs on Twitter.* Knight Foundation.

Koch, Janin. 2017. "Design Impliations for Designing with Collaborative Ai." *AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems.* AAAI Spring Symposium. 415-418.

Lievens, Eva. 2019. *A human rights perspective on our algorithmic transparency.* Ghent: Human Rights Centre Ghent University.

Minsky, Marvin. 2007. *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind.* Simon and Schuster.

Weizenbaum, Joseph. 1967. "Contextual Understanding by Computers." *Computational Linguistics* 10 (8): 474-480.

Yang, Qian, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steifield. 2018. "Investigating how experienced UX Designers Work with Machine Learning." *DIS 2018.* Hong Kong: DIS. 585-596.