

Moderating the self

Digital cultures and politics connect people even as they simultaneously place them under surveillance and allow their lives to be mined for advertising. The collection and the analysis of data therefore, becomes a necessity. And as data becomes a valuable source of information, mechanisms of collecting and moderating them are needed.

The act of moderation though is hard to examine, as it's intentionally designed to be easily overlooked, it is meant to disappear, but it does so for some more than others.

On this idea of the intentionally disappearing design, Geert Lovink and Mieke Gerritze argue in *Design Manifesto* "When everything is destined to be designed, design disappears into the everyday. We simply don't see it anymore because it's everywhere. This is the vanishing act of design."²¹ In the same way Wendy Hui Kyong Chun refers to the habitation of social media pointing out the seemingly naturalised way that technology has been embedded into our lives, "technology matters when it seems it doesn't matter at all"²²

Platforms likewise, they are omnipresent, embedded in the everyday life of the user, as they're designed to "disappear" and be seemingly unnoticeable. "The platform would like to fall away, become invisible beneath the rewarding social contact, the exciting content, the palpable sense of community."²³

Content moderation is partly what distinguishes them from the open web: they moderate (removal, filtering, suspension), they recommend (news feeds, trending lists, personalised suggestions), and they curate (featured content, front-page offerings).

Most platforms divide their rules by providing two distinctive documents. The "terms of service" is basically a contract that explains, in a more formal tone, the terms under which user and platform interact, the obligations users must accept as a condition of their participation. The "community guidelines", focused on the proper use of the

platform, is the one more possible to be read, most likely when users find themselves facing content or users that

offend them. What is immediately observable is the distinctly casual tone of the community guidelines, compared to the denser legal language of the terms of service.

Moderation is what shapes social media platforms as tools, as institutions, and as cultural phenomena.²⁴ This set of rules and procedures have consolidated into functioning technical and institutional systems, that sometimes manage to disappear, and sometimes create friction between the platform and the user.

In "Critical Mass", Jonas Lund criticises the relation between platform and user, posing the question of agency in social media. The artist created a space filled with objects accordingly to the audience's choices. Using his own words "To paraphrase Agent Smith from the Matrix – what good is a voice if you are unable to be heard?"²⁵

Social media content, hashtags etc are censored by automated systems, according to the platform's community guidelines. It is known how certain bodies and perspectives are being policed on Instagram, and how certain people are targeted and silenced, by conscious or unconscious bias built into the framework of our digital world. In this respect, the online newspaper Salty distributed a survey on to its followers on Instagram and via newsletter. As a community for and by women, nonbinary and queer people, the Salty's group of followers is heavily composed of these demographic groups. As such, the survey surfaced issues that are affecting these communities. The report established some findings. Queer and women of color are policed at a higher rate than the general population, respondents in the survey who were unable to advertise their products were more likely to identify as a cis woman than any other identity type; plus-sized and body-positive profiles were often flagged for "sexual solicitation" or "excessive nudity"; policies are meant to protect users from racist or sexist behaviour are harming the very groups that

²¹ Gerritzen, M. and Lovink, G., 2019. *Amsterdam Design Manifesto*. The Image Society, p.7.

²² Chun, W., 2013. *Habitual New Media*.

²³ Gillespie, T., 2018. *Custodians Of The Internet*. Yale University Press. p.7

²⁴ Ibid. , p.6

²⁵ Jonas Lund. 2020. *Jonas Lund - Critical Mass*. [online] Available at: <<https://jonaslund.biz/works/critical-mass/>> [Accessed 19 June 2020].

need protection. For example, people under attack for their identity have been reported or banned instead of the attacker.²⁶

Elaborating on moderation and biases, algorithms are the backbone of content moderation. Algorithmic models produce probability scores that assess whether the user-generated content abides by the platform's community guidelines. An algorithm is simply, a process or set of rules usually expressed in algebraic notation.²⁷ Thus, an algorithm, as a notion of mathematics, can be interpreted as a language of precision, calculability and predictability. Therefore, we could say that an algorithm radiates an aura of computer-confirmed objectivity.

Algorithmic bias is found across platforms including search engines and social media platforms. These platforms collect their data from online user behaviour, creating significant amount of data every day. Therefore, as human behaviour contains value and judgement, online behaviour will show something similar. This online behaviour will express itself in data from which the algorithms learn, as data is a mirror of human behaviour, and an algorithm acts as the reinforcer. The data and the programmer are the factors that transmit the judgment and subjectivity on the algorithm. Therefore, Blaming 'the algorithm' or 'big data' is like turning a blind eye : we fail to see the real source of malice.

However, developments in machine learning have enabled algorithms to self-optimize and generate their own improvements. They can now self-author and self-create. Recognising patterns allows humans to predict and expect what is coming, which also applies to self-learning algorithms. This greatly complicates notions of authorship, agency and even algorithms' status as tools. The knowledge that a self-learning algorithm can possess through the collection and optimisation of data, is actually a reflection of our knowledge, as it is the creator who actually gives the algorithm a world to learn from. We could perceive algorithms as a tool of self-reflection, as we embed our values and ideas into algorithms, only for them to reveal what we know about ourselves.

Donna Haraway, has an interesting take on the way that high-tech culture challenges this dualism of human and

machine. "It is not clear who makes and who is made in the relation between human and machine. It is not clear what is mind and what body in machines that resolve into coding practices...Biological organisms have become biotic systems, communications devices like others. There is no fundamental, ontological separation in our formal

knowledge of machine and organism, or technical and organic."²⁸

What is important to keep from this, is that the blurred notion of agency is not explicitly applied in algorithmic processes, but technologies in general, as they do not, in themselves, change anything, but rather are socially constructed and deployed. That means that the possibilities of a new technology take shape in the hands of those with the greatest economic power.

²⁶ Salty. 2020. *Exclusive: An Investigation Into Algorithmic Bias In Content Policing On Instagram*. [online] Available at: <<https://saltyworld.net/algorithmicbiasreport-2/>>.

²⁷ Uricchio, W., 2019. *Situated In Translations: Cultural Communities And Media Practices.*, Transcript, Bielefeld,p.32.

²⁸ Haraway, D., 1991, *Simians, Cyborgs and Women : The Reinvention of Nature*, New York; Routledge, pp.177-178