

Datafication unmasked 1.

“AI, ain’t I a woman?” seems an odd question to pose to Artificial Intelligence. How would technology that is learning how to create itself misgender someone – and especially Michelle Obama? In a passionate poem, poet of code and computer scientist Joy Buolamwini (2018) shows how the algorithms that enable the most used facial recognition systems often misgender women of colour, bringing the examples of iconic women being wrongly matched to the male gender. Not only does she unveil the fallacies of the technology, but she exposes the harm that they implicitly inflict. As a black woman herself, she struggles to find representation in the most advanced technologies. Her frustration goes further: while conducting research at MIT, facial recognition software could not detect her face, until she wore a white mask.



Joy Buolamwini, image retrieved from [Insurance Journal.com](https://www.insurancejournal.com)

Joy Buolamwini’s story is one among many. In Michigan, Robert Julian-Borchak Williams faced an arrest, following a faulty match from the facial recognition system employed by the police (Hill, 2020). In the US, millions of black people were denied access to healthcare programs, as the algorithms used to allocate them continuously advantaged white people (Ledford, 2019). These stories of discrimination are increasingly populating news headlines, gaining resonance in the public discussion. Meanwhile, scholars are warning over the exploitative practices led by big tech companies through the endless capture of user generated data (i.e., Noble, 2018; Zuboff, 2019), exposing the mystified character of the narrative surrounding big data and digital technologies.

In this context, data activist groups are arising to oppose data mining and algorithmic bias, as identities online are increasingly datafied.

But what is datafication?

The term “datafication” is used to define the transformation of social action into “online quantified data” (van Dijck, 2014, p.198), enabling for real-time tracking of human behaviour and its predictive analysis (ibid.). While creating an extremely vast database that allows for unprecedented research opportunities, datafication has become a legitimate way to access and monitor online behaviour (ibid.), raising questions and criticism over the dynamics of power this practice entails. Researches bring under scrutiny the very epistemological and ideological premises of datafication (ibid.; Couldry, 2014), as well as its biopolitical and commodifying foundations.

The underlying ideology of datafication is dataism, a universal narrative that proclaims algorithms and big data as legitimate forms of authority (Harari, 2016). Dataism and datafication are built upon the belief in the objective quantification of human behaviour through digital technologies, requiring trusting on the data-collecting agents and institutions (van Dijck, 2014), and to endorse logical positivism as the philosophical assumption and approach toward reality.

Seeing the world and human activity as entirely and perfectly knowable, this ideology sustains and promotes an all-encompassing quantification process that admits no bias (Greenfield, 2017). Data is perceived as transparent and uncompromised, a transferable knowledge that responds to the quest to gather the many quantified selves available online through the instrumentalization of their body and the capture and monitor of their behaviour (ibid.).

This paradigm implies two fundamental assumptions. On the one hand, that extracted data are spontaneous traces left in neutral technological channels, emptying them of much of the algorithms and business models that structure the specific technological mediation they enable (van Dijck, 2014). Secondly, the supporters of datafication see the relationship between data and people as straightforward and self-evident, sustaining the idea that data can explain and predict people’s actual behaviour in disregard of the interpretative framework that is at work in every attempt of explaining patterns through data analysis (ibid.).

We can refer to it as “the myth of big data”, a contemporary form of cultural storytelling. A myth that sustains the usage and manipulation of data as *terra nullius* - land of no one. More about it next month.

Consulted resources

Buolamwini, J. (2018). *AI, Ain't I A Woman?* [video].
<https://www.youtube.com/watch?v=QxuyfWoVV98>

Couldry, N. (2014). A Necessary Disenchantment: Myth, Agency and Injustice in a Digital World. *The Sociological Review*, 4(62), 880-897.
<http://journals.sagepub.com/doi/10.1111/1467-954X.12158>

Greenfield, A. (2017). *Radical Technologies: The Design of Everyday Life*. London and Brooklyn: Verso.

Harari, N. H. (2016, August 26). Yuvan Noah Harari on big data, Google, and the end of free will. *Financial Times*.

Hill, K. (2020, June 24). Wrongfully Accused by an Algorithm. *The New York Times*.

Ledford, H. (2019, October 24). Millions of Black People Affected By Racial Bias in Health-Care Algorithms. *Nature*. <https://www.nature.com/articles/d41586-019-03228-6>

Noble, U. S. (2018). *Algorithms of Oppression. How Search Engines Reinforce Racism*. New York: New York University Press.

van Dijck, J. (2014). Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Secular Belief. *Surveillance & Society*, 12(2), 197-208.
<http://library.queensu.ca/ojs/index.php/surveillance-and-society/article/view/datafication>

Zuboff, S. (2019). *The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.