

The thirst for data generated by digital slaves

Artificial Intelligence and Democracy: Risks and Promises of AI-Mediated Citizen-Government Relations is a book by Jérôme Duberry, published in 2022 as the result of a research project funded by the Swiss National Science Foundation.

The book structurally and thematically spans across several important issues which AI as a digital technology brings into the citizen-government relations, methodologically relying on literature and policy review, workshops, focus groups and semi-structured interviews with authorities in citizen participation, online platforms and artificial intelligence. It begins with the conceptualization of AI and ends with the ramifications of civic tech as the technology used to increase citizens participation and civic engagement. In between, it explores the most common positive and negative effects which algorithms, automation and big data exert over citizen participation in policy making. In fact, given that the participation of citizens in policy-making is a direct expression of popular sovereignty and this participation is based on the assumption that citizens are (1) well informed (2) consulted and (3) included in the decision process, the author presents in separate chapters how the use of AI contributes to these assumptions. The findings, in brief, show that AI creates filter bubbles and echo chambers alongside the digital space allowed for people's opinions and initiatives; AI assists in 'digital listening' and mass-surveillance of these practices not only for national security purposes, but for pure profit; and finally, it also capitalizes *via* digital advertising on people's online presence, particularly by gearing their opinion with (individually optimized) propaganda.

A separate chapter is dedicated to the global context in which information is often weaponized to influence the (geo)political landscape. The chapter is limited only to the propaganda used by one country due to its 'long history of information weaponization' (cf p.164), and thereof the greater accessibility of the sources for analysis. We will leave it to the imagination and personal biases of the reader to guess the particular country at question whose operations abroad seek to reduce, so goes the framing narrative, the influence of liberal democracies and democratic values.

Rhetorically, the style of writing is restrained and agreeable rather than polemical in its almost impeccable juxtaposition of various referenced opinions. The book actually abounds

in well-referenced considerations from multidisciplinary literature and relies heavily on institutional and international policy documents, recommendations and definitions. At points, this can cause and propagate a sense of losing the wood for the trees in the argument line (near to the impression one is left with after having read an EU-research policy document as such). The book is accompanied with an index of terms and names, simplifying the reader's exploration of the topic. Brief insight is offered about the different 'traps' and failures to account for the interactions between technical systems and social worlds, about the voluntary ethical codes and guidelines developed around the world with regard to AI governance or the principles that should guide the adoption of AI. Handy AI taxonomies that take into account conceptual and research perspectives have been adapted from different sources and presented. The reader will also get a well-rounded presentation of the policy process given the accent the book places on citizens' participation in policy-making. And while the task the author has taken upon himself is both commendable and timely, particularly given the relevance of the question of AI to the democratic project in terms of who is using this technology, for what purpose, how does the use of AI influence the trust of citizens in the democratic institutions and power relations in policy-making, the philosophical questioning within which it sets to answer these questions on occasions remains sketchy.

For instance, the author chooses to consider technology from the so called *(co-)evolutionary innovation studies* perspective, which basically claims that tech is simultaneously 'a tool and an outcome' of governance, a favourite and comfortable trope of the social science scholars of our time. In addition, the author considers the conceptual challenges of AI, starting from the Dartmouth Summer Research Project in 1956 and the historical change in its development from *symbolic* or *conceptual AI* in the period where researchers ambitiously aimed at configuring the ways to make machines use symbols, through the period of decline and disinterest in AI due to the failure to achieve that goal and the resulting disenchantment, until the last decade of the 20th century when computing power and data availability began to rise and *statistical AI* became reality. Now, even though AI is a set of algorithms geared operationally towards problem-solving, the author aims to define it less from a technical and more from a 'social science perspective' which for some reason seems to go hand in hand with *substantiating* AI, which is to say referencing algorithms as though AI is capable of observing **its** environment (cf. p21, my emphasis). What is questionable in this definitional concept is not so much the 'relational/agential' choice of words, but the

very fact that such approach exempts *non-reflectively* the key agents – the human species – from the dynamical political trajectory they (might) create with their governments. Put differently, the risks and promises of AI (as an intermediary in the explored relation) are primarily and ultimately the risks and promises of the *political* will of those who create, select, implement, dismiss or favour one over other set of algorithms. This claim is perhaps ‘naïve’ from the systems theory philosophical framework overarching the co-evolutionary innovation studies perspective to which sides the author. Yet, by leaving untreated this theoretical space under which Duberry wants to subscribe his own research, he has missed the opportunity to challenge the reader along a bolder critical (critical in the sense of gr. κρίνεις, i.e. a discerning) line of thought that could set the tone for an in-depth treatment of his chosen topic. Epistemologically, we can never overestimate humans precisely because of their capacity for meta-representation of the lower-level constraints they use (which is actually their *differentia specifica* with regard to artificial ‘intelligence’). But politically, the awareness of the *governing* effect of software on social behaviour is far from sufficient. Moreover, not underestimating technology might quickly slip towards our bowing to yet another golden calf.

And we are left wondering, as digital slaves in solitude; how are we to counteract the imposed political constraints and *risks* rather than promises of the AI-mediated reality and citizen-government relations? If technology for Heidegger revealed the world as raw material, the case of AI reveals even more so emphatically the *human subject* as raw material, available for production and manipulation. Is the civil society autonomous and able to resist manipulation by the state and business interests? Are the elite committed to democratic values such as popular sovereignty? Are they committed to employing all possible means to decrease the massive asymmetry of power between citizens and governments, rather than increase it? Do the political, business and intellectual elites value the lives of others as *indispensable* and sacred? If the answer is negative, then the political imaginary of liberal democracies will continue to erode, weighed by the hypocritical abyss and the non-coincidence of our proclaimed values vs. demonstrated social practices of heads of state, CEO’s, politicians, organizational leaders etc. Moreover, any further treatment of the role of AI as a political agent under, say, the co-evolutionary framework of emergence, non-linear dynamic etc. *without treating properly the conditions* necessary to effectuate change would be a lacklustre (even if fancy) avoidance of the deep muddy waters of any veritable political questioning.

This is not to say that the above mentioned non-coincidence will necessarily provoke a slip towards authoritarianism, as contradictions are far from rare in democracies (just consider our current decreased social cohesion due to less shared experiences, personalization of constructed realities, big data sets under exclusive control of global companies, individual profit maximization with no social responsibility, weakening of public-interest goals and social ordering 'designed' by global IT companies effectuating societies dislodged from their national context), but what is certain is that if sovereignty of people continues to be deconsolidated as a major reference of the democratic imaginary, a new form of society will be created, one in which every questioning of coercion will be dismissed as malevolent or insane.