

## AI technologies and democracy: concerns for today and tomorrow

Presentation at expert workshop on ‘Building Resilient Democracies? The Global Significance of EU Initiatives to Safeguard Digital Citizen Engagement and Online Democratic Processes’, at the Graduate Institute of International and Development Studies, 9 December 2019

Abhimanyu George Jain

My research on artificial intelligence or AI is usually in the context of military applications of AI, in the context of the SNF-funded project on ‘LAWS & War Crimes’, led by Professor Paola Gaeta. That project is concerned with criminal responsibility for breaches of international humanitarian law by human-supervised lethal autonomous weapons and AI-supported human soldiers.

Today, however, I am going to be speaking about AI technologies and democracy. As is the case with digital technologies, AI technologies can also be used in support or furtherance of democracy. These technologies can be used to mitigate threats to democracy, for instance, by detecting fake news. Or they can be used to make governance processes more responsive and fact-based, as in algorithmic regulation, where governance is driven by algorithms and big data analytics. Or they can be used to supplement or even replace institutions which are integral to democratic systems, such as the judiciary. If that last one seems a little far-fetched, it’s worth highlighting here that the UK and Estonia are both engaged in developing and testing AI-driven processes for small civil claims, and some U.S. jurisdictions already use algorithms in making decisions about eligibility for bail.

In my presentation today I want to talk about these present and future uses of AI in support of democracy and I want to set out some concerns that I have in this regard. I have four broad concerns.

The first is the possibility, to paraphrase David Kilcullen and Andrew Exum, that our turn to AI (and perhaps even to digital technologies) represents a technology substituting for a strategy.<sup>1</sup>

Kilcullen and Exum devised that formulation in the context of American military uses of drones for counter-insurgency operations in Pakistan, so let me start with an example from the military context to explain what I mean here. A key concern in ‘new’ wars is identifying who is a civilian and who is a combatant and consequently, who can be attacked. In armed conflicts with groups like ISIS, whose members are not distinguishable by their uniforms, difficulties in identification and targeting are seen as a major source of avoidable civilian harm. In this context, drones and lethal autonomous weapons have been advertised and embraced on the basis of the enhanced precision that they offer – “*putting warheads on foreheads*”. The problem with this narrative is that while there certainly is enhanced accuracy in targeting, there is no corresponding enhancement in the identification that precedes the targeting process.<sup>2</sup> These technologies make it possible to kill a single individual in a crowd, but do not significantly

---

<sup>1</sup> David Kilcullen and Andrew McDonald Exum, ‘Death From Above, Outrage Down Below’ *New York Times* (9 May 2009) <<https://www.nytimes.com/2009/05/17/opinion/17exum.html>> accessed 5 December 2019.

<sup>2</sup> In this regard, see generally, Grégoire Chamayou, *Drone Theory* (Janet Lloyd tr, 2015).

enhance the ability to determine if that person was a combatant. The reason for this is that humans themselves do not know precisely how to do that in the context of contemporary armed conflicts. A soldier who would not be able to identify a Taliban fighter if she passed him on the street, cannot necessarily identify him through the grainy infrared feed of a drone, and nor can she describe that process for coding in an algorithm.

In other words, we cannot expect our technologies to achieve results that we cannot teach them to achieve. Using that as a springboard, let me turn to the use of AI to detect ‘fake news’. Putting aside the issue of deliberate disinformation, the vast majority of ‘fake news’ distorts frames and perspectives rather than facts. Correcting those distortions is a subjective process that involves the balancing of a large number of competing interests, including individual rights such as freedom of expression, that are themselves essential in a democracy. Reasonable humans can disagree on these subjective assessments: the idea that we can reach consensus on how to train a machine to do so demands careful scrutiny.

More importantly, one must pause here to wonder whether the scrutiny of frames and perspectives really is a feasible or productive endeavour? If one takes the view that all information is mediated by frame and perspective, and that objective truth is elusive if not non-existent, then this endeavour itself is questionable.

To sum up here, my first concern is that our turn to technology in this moment of seeming crisis represents an unreflective search for technocratic solutions to problems that we haven’t entirely understood, and in doing so we risk using technologies as strategies.

My second concern relates to the idea that our technology shapes us just as much as we shape our technology.

Donna Haraway wrote of ‘our cyborg selves’ in 1985,<sup>3</sup> well before the advent of the current information technology age. The concern that the inter-twining of our lives with technology would change us was very real then and is very real now. But the advent of algorithmic regulation signals a remarkable qualitative jump. As Eyal Benvenisti has noted, the use of AI in governance is revolutionary in its potential to eliminate actual communication between the governors and the governed.<sup>4</sup> Across the many significant technological changes that have come before this, the centrality of this bilateral channel has remained unaffected. Now this is under threat.

To explain the nature of this threat, let me use an example. I may choose to spend every Friday night at the movies instead of going to the opera. But this does not mean that I would support the de-funding or abolition of the opera. This may seem like an outlandish and exaggerated example, but it illustrates the problem with the elimination of communication between the governors and the governed. It eliminates the crucial distinction between what we do and what we want to do, or more importantly, between who we are and who we want to be. And this foreclosure of possibility in itself represents a fundamental change in who we are.

Let me try and explain this further with a more concrete example. The closer our society moves towards an ‘algocracy’, the more difficult it becomes for the individual to stay off the grid and

---

<sup>3</sup> Donna Jeanne Haraway, ‘A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century’, *Manifestly Haraway* (University of Minnesota Press 2016).

<sup>4</sup> Eyal Benvenisti, ‘Toward Algorithmic Checks and Balances: A Rejoinder’ (2018) 29 *European Journal of International Law* 1087.

not become a data object. If the individual wishes to retain political agency, they must be visible to the algorithm. The act of political participation changes from casting votes to accepting cookies; the consequences of that act change from the selection of a human-intermediated government to the shaping of all-pervasive regulation; the meaning and scope of political protest changes; indeed, the very concept of the political will change. In this regard, it is interesting to note that much of the EU policy literature on AI speaks of the importance of literacy and leaving nobody behind, but references to opt-out rights are far fewer in number.

My third concern, paradoxically, relates to the fear that AI might prevent change and lead to ossification. Some part of the AI ecosystem will inevitably be static: the way data is collected, or the way data is characterised, or perhaps even the original question or task put to the algorithm. To the extent that that is the case, there is the possibility of some form of human bias afflicting that static element. In and of itself this would not be such a problem; after all, the very same biases afflict human decision-making. The problem, however, is that the discursive construction of AI, emphasising its objectivity and its dynamism, increases the risk that the bias remains unconsidered, unnoticed and unaltered, forever binding human society and politics to a 21<sup>st</sup> century idiosyncrasy.

Now, I am very willing to concede that I am not a technical expert, and perhaps the scenario which I have just sketched is preposterously wrong. Nonetheless, it does point us to an important consideration. A crucial element of the AI discourse is its objectivity and omniscience. Mark Andrejevic describes this as “*a fantasy of a truth unpolluted by any deliberate human intervention – the ultimate in scientific objectivity*”.<sup>5</sup> Returning to a point I made briefly in the context of my first concern, it is hard to defend the idea that there is any such thing as an objective truth. A few years ago, I met somebody who was pursuing a PhD in physics at Oxford, and to my incredible confusion, he explained his doctoral project as an effort to devise a more accurate scale to measure a metre. Even a seemingly objective measure of distance cannot fulfil the ‘fantasy of a truth unpolluted by human intervention’.

Returning to the idea of change and ossification in the context of AI, if the ‘objective truths’ that algorithmic regulation is producing are themselves contingent and transient, then we must beware the possibility that the turn to technology impedes human evolution.

A fourth concern relates to the effect that the adoption of these technologies in the developed states of the west has for less developed states. One aspect of this concern is that the rush to adopt these technologies may itself contravene democratic processes. An important aspect of the EU policy documents on AI is the commitment to development and deployment of AI in consonance with European principles, particularly European data protection regulation. This commitment recognises that mainstreaming AI without privacy controls contravenes democratic values. Yet, there is a very high risk that that might well happen in the global South. Consider, for instance, the fact that a relatively strong democracy like India, which has an ambitious AI policy, doesn’t have a data protection law.

This brings me near the end of my presentation. To summarise, I started by introducing the idea that AI technologies, like digital technologies, could be used in furtherance or support of democracy. And then I set out four concerns about these present and future uses of AI: first, that our turn to technology feels suspiciously like a technology masquerading as a strategy;

---

<sup>5</sup> Cited in William Davies, ‘Why Can’t We Agree on What’s True Any More?’ *The Guardian* (19 September 2019) <<https://www.theguardian.com/media/2019/sep/19/why-cant-we-agree-on-whats-true-anymore>> accessed 5 December 2019.

second, that our turn to technology will inevitably change us; third, at the same time, the turn to technology may prevent us from changing; and, fourth, that the adoption of these technologies in the global North will have ripple effects in the global South.

I should clarify that this is not an exhaustive list of concerns regarding the use of AI to strengthen democracy. Nor is this a very original list; these concerns are all reflected to varying degrees in existing policy and academic literature. I have chosen these four concerns to speak about today, because I thought that they were of particular interest for our discussion today, and because these four concerns are relatively under-explored in the policy literature.

Before concluding, I would like to make three final observations to contextualise the four concerns that I have just set out.

First, at this point I must sound entirely like a Luddite. I will confess that I am extremely sceptical of these technologies. But I do believe that these technologies present important benefits, and that in the coincidence of costs and benefits, these specific technologies are no different to prior technological upheavals.

This brings me, however, to my second point – the relative proportion of costs and balances. In looking through the European Commission’s 2018 document on the European perspective on AI,<sup>6</sup> I came away with the distinct impression that while there are many, definite and significant costs of AI, the benefits are frequently speculative or less significant. This certainly corresponds to my understanding of the contingent and hypothetical promises of precision that dominate the discourse on lethal autonomy. If this characterisation of the relative costs and benefits is true, then we must beware the technological somnambulism or false necessity that is leading inexorably to these digital and algorithmic futures.

My third and final point relates to the broader scope of today’s discussion. A key premise seems to be idea of threats to democracy against which democracy must be made resilient. This shines through particularly strongly in some of the EC documents on digital democracy and AI. If the threat in question is Russian hackers, that is entirely reasonable. But if the threat is right-wing populism, then I have my doubts. Not only does that seem definitionally undemocratic, it seems to endorse the idea of a liberal objectivity. This is the same liberal objectivity which supports the veneration of algocracy and which I have tried to argue against today.

---

<sup>6</sup> European Commission and Joint Research Centre, ‘Artificial Intelligence: A European Perspective’ (2018) <[http://publications.europa.eu/publication/manifestation\\_identifiers/PUB\\_KJNA29425ENE](http://publications.europa.eu/publication/manifestation_identifiers/PUB_KJNA29425ENE)> accessed 6 December 2019.