



Editorial: Beyond regulatory ethics

Satinder P. Gill¹

Accepted: 30 March 2023 / Published online: 20 April 2023
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2023

The recent news headline of a petition (O' Brian 1), with names such as Elon Musk, Steve Wozniak, and Stuart Russell, calling on all AI labs to halt work for 6 months on the training of AI systems more powerful than GPT-4, speaks of an “out-of-control race to develop and deploy ever more powerful digital minds that no one—not even their creators—can understand, predict, or reliably control” with “catastrophic” risks of disinformation and automation. The petition is timely for this special issue on AI for People. This out-of-control race has tended to prioritize narrow and short-term economic interests over broader and longer-term societal needs, and the space race narrative is a visible example of this. In this special issue, authors are questioning underlying assumptions and proposing design methods to mitigate issues such as bias, and envisioning future skills requirements for interdisciplinary designers, and ways of bringing together stakeholders across society for creating AI for social good.

Where the emphasis on bias is on technical aspects, could our diverse experiences of bias in daily life provide an alternative way to address the problem of bias in the machine? One recurring theme in this issue is the gap between the generation and application of generic and abstract rules and the level of an individual's context/situation. A limitation of current big data analytics is being able to have an understanding of the differences and diversity at the level of the individual, for example, the uniqueness of a patient and their condition. There is a need for ethical concerns with justice, fairness, and bias with regards to AI to be addressed at both a generic and individual contextual levels. If we desire technology for social good, it needs to have the potential to adapt empathically to individual human lives.

As we engage more and more with the artificial representations of emotions embedded in machines this may have

an impact on how we engage with other humans and handle our own emotions. How is emotion expressed, felt and acted upon in our lives with others beyond the explicit facial expression or movement of a body part of a machine? Perhaps our quest for AI can shed more light on what it means to be human.

Can we cultivate interdisciplinary design cultures that can create an AI that can be managed in an ethical way for social good? Could we, for example, educate future philosopher engineers to handle the breadth and depth of ethical and social concerns. The recent petition is a top down call in the name of public interest. However, in order for AI to actually serve a public interest perhaps we need to shift the focus of the discourse towards participatory and inclusive processes when developing AI systems. Such an approach could alleviate the marginalization of those who are vulnerable and at the receiving end of AI automated decisions.

We are also more than individuals. We are persons, and at any moment we may be a child, a parent, a patient, a clinician, a colleague, depending on whom we are engaging with. Furthermore, we exist and live in cultures and communities that are complex and richly diverse. Bias, ethics, care, decisions, judgements, are not made outside of our cultural practices, yet in the out-of-control race, machines are designed to regulate and control decisions, ethics, and behavior. Within social robotics there is a growing recognition that culture needs to be considered in order for trust to be fostered. However this needs to go beyond a normative construct of culture and diversity, as if these were static and not evolving.

The petition raises the question of what is a meaningful data science, and for whom and why? Interdisciplinary data science should be as much a methodological process as a political and interpersonal process. Rather than focusing on regulatory frameworks to manage AI, this special issue is a ray of hope where interdisciplinary researchers are questioning and coming together to provide humanistic solutions.

✉ Satinder P. Gill
spg12@cam.ac.uk

¹ Centre for Music and Science, University of Cambridge, Cambridge, England

Reference

Matt O'Brian (2023) Musk, scientists call for halt to AI race sparked by ChatGPT. <https://apnews.com/article/artificial-intelligence-chatgpt-risks-petition-elon-musk-steve-wozniak-534f0298d6304687ed080a5119a69962> (Accessed 30 Mar 2023)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.