OPINION PAPER

The ethics of technology innovation: a double-edged sword?

David De Cremer[1] · Garry Kasparov[2]

Abstract

We discuss the dilemma that while AI is considered as one of the most powerful engines that exist today to drive innovation, at the same time, however, rapid application of AI also has the potential to further increase inequalities and societal harm. This makes that we are being confronted with the question whether today's amazing tech innovations may ultimately not bring limited benefits to the weaker members of society who need this kind of innovation the most. As such, do we need to slow down tech innovation to ensure that not more (and possibly new) unethical outcomes emerge over time? We note that the pursuit of tech innovation has been advocated primarily to optimize productivity, and, hence, economic growth. This pursuit represents, however, a narrow perspective on the good that AI could produce. Therefore, we argue that we need to adopt a less narrow perspective on what optimization means using tech innovation in ways that optimizes a diversity of human interests. By adopting such broader perspective, we propose an integrative approach where we start from the idea that we need to continue pushing tech innovation, but in combination with regulating innovation efforts and installing a stronger sense of moral awareness and responsibility among those in charge of the tech innovation journey. We conclude with outlining recommendations that can help promote this integrative approach, including the combination of self- and government regulation, promoting training efforts to establish more responsible leadership, and encouraging efforts to bring AI faster to the people.

Keywords Technology innovation · Inequalities · Big tech companies · Responsible leadership · Artificial intelligence · Self-regulation · AI governance

It is an understatement to say that because of technology, our society is undergoing massive changes. No one can escape from the tendency of governments, companies, and society to place an emphasis on gaining digital skills. Indeed, the introduction of intelligent technologies in all domains of our lives is something we all consider as inevitable by now [1]. The expected economic benefits are significant—with estimates of boosting the global economy by around US$16 trillion by 2030 [2]. All of this makes that Artificial intelligence (AI) and all its subcategories (e.g., Machine Learning, Robotics, Vision, Natural Language Processing) are recognized as the next necessary step in our drive for innovation. In fact, AI is judged as an even more important disruptor than the internet once was [3].

Dealing with the inevitability of AI adoption, it follows that data science and AI courses are gaining in popularity. There are even coding courses for pre-schoolers to introduce them already to a future that will be digital. Clearly, we are getting ready—and are being prepared—for a world where AI will impact almost everything we do. Having said this, while AI gains importance in decision-making in various aspects of our lives, we also observe that there is an increasingly greater scrutiny on whether AI is "ethical" and will benefit humans in ways that adheres to the normative expectations of justice, morality, and trust. Will the AI system that helps to filter job resumes be biased against a minority race, because previous position holders were not from that race? Will facial recognition technology lead to the wrongful arrest of some people? Will the algorithm recommend certain beauty products to the consumer because they bring in higher profits for the company?

The recognition that AI adoption may not reveal outcomes that are neutral, fair, and appropriate as we initially believed (i.e., AI as a rational decision-maker should be expected to be unbiased), makes that as a society, we are

✉ David De Cremer
  bizddc@nus.edu.sg

1 Centre on AI Technology for Humankind (AiTH), NUS Business School, National University of Singapore, 15 Kent Ridge Drive, Singapore 119245, Singapore

2 Renew Democracy Initiative (RDI), New York, NY, USA

being faced with an intriguing yet complicated dilemma. The dilemma is that AI is very much embraced as one of the most powerful—if not, the most powerful—engine to drive innovation. At the same time, however, the fast and relentless application of AI seems to sustain and potentially even further increase inequalities, societal harm, and threats to values that we consider—from an ethical point of view—necessary to be present in our lives. We have started, for example, to wonder how we deal with the fact that we are willingly conceding so much personal data out of convenience, and who benefits most from our willingness to do so? Because people want to use all these benefits generated by AI, they do not always recognize the simple truth that by being digitally connected their information is collected by known and unknown parties. This kind of connection makes them vulnerable to exploitation and possible unfair treatment. So, we are being confronted with the question whether the amazing tech innovations, that we are witnessing and implementing today, may at the same time not bring too many dangers for society to the fore, and especially so to its weaker members? This is an impactful question, because if we would arrive at the conclusion that tech innovation is promoting the interests of a few, but not of the many, what should then be our most ethical response? In that case, may it then be that we need to slow down tech innovation to ensure that not more (and new) unethical outcomes emerge over time?

To be clear from the start, we are pro tech innovation. It cannot be denied that intelligent technologies have the potential to enrich our lives in all kinds of *different* ways. However, when we talk today about tech innovation, we seem to be focusing only on enriching society in *one* specific way, which is increasing employees' productivity to drive economic growth. This rather narrow view may not be so surprising. After all, wasn't it noble price winner in economics Paul Krugman, who famously said, "Productivity isn't everything, but in the long run it is almost everything." Why? Innovation is seen as the engine to grow productivity and in that process, it is the workers who contribute input to this system. Obviously, the more input they provide, the more output will be achieved, so productivity can keep growing. Humans, however, have physical and mental limits. Therefore, at one point in time, the abilities of the group of people striving for increasing productivity levels will stop growing as they will hit their limits. Fortunately, with the arrival of intelligent technologies, we have found new ways for the same number of people to keep growing in productivity by augmenting their abilities via technological advancements. As such, the fast—and perhaps not so reflective—advances in AI systems are needed to keep productivity levels rising and innovation growing [4].

This is, of course, not a new idea. The British economist Johan Maynard Keynes predicted almost a century ago that tech innovations would improve productivity in labour [5]. And, indeed, this is happening now. Keynes, however, also predicted something else. He also assumed that growth in productivity—because of tech innovation—would lead to people being paid higher salaries and having to work less. However, this prediction is not materializing. People today are still working long hours, salary increases are rather exception than default, and pension funds are under threat forcing people to work longer than before. At the same time, however, wealth is accumulated in the hands of a small group of people, notably in charge of the companies developing the technology that we are adopting today. Take, for example, the message of Jeff Bezos after he successfully made a tour of space in his Blue Origin spacecraft. Upon his return to earth, he thanked the Amazon employees and customers, because they paid for his trip. Not surprisingly, people could not appreciate his message [6].

The pursuit of tech innovation has been advocated to optimize productivity, and, hence, economic growth, and therefore those with more economic power are destined to gain more than others. So, if we want more people to benefit from today's technological advancements, we will need to adopt a less narrow perspective on what optimization means. Specifically, we will need to use tech innovation in ways that optimizes a diversity of human interests. As a matter of fact, we believe that the dogma to see technological advancement as the primary way to optimize human performance does not adequately clarify what optimization is about and what benefits it will bring to humanity. It is our opinion that discussions are lacking about the shape that human-centred approaches to tech innovation should look like, because surely those efforts must be broader than only optimizing the economic prospect of growth and profit.

As Shannon Vallor notes in her excellent blog, "Optimising is defined in the value that we strive for. If we say optimizing is to make people more productive so they can keep innovation growing, then we are pursuing an efficient system where being human and serving human interests and diversity is not the end point" [7]. An ethical point of view, however, requires that we advocate a human-centred approach that serves a diverse set of humans needs and concerns. The result of such approach will be that it will inevitably reveal both optimal and non-optimal outcomes, depending on which perspective you take. For example, for some people, AI advancements need to be evaluated whether it contributes primarily to people's mental well-being and thus becomes more nurturing rather than fast-moving, whereas for others, this kind of application may be regarded as an undesirable outcome because it reduces the economic profit one could make if AI adoption continued in a fast pace. So, if one wants to adopt a fairer and just approach to looking at tech innovation to move forward as a society, a more diverse way

of evaluating what technology advancements can and cannot optimize will be needed.

The truth, however, is that tech innovation to promote optimal outcomes in today's economic world is and remains defined in narrow ways. And, for this reason, the relentless pace of AI adoption and employment is expected not to slow down. Indeed, a dominant capitalist system calls for ever-increasing growth and thus for an increase in speed rather than a slowdown. So, our reflective question whether tech innovation should be slowed down to safeguard our ethical compass may in such a system be totally irrelevant. In fact, the default thinking is that with much-needed tech innovation, there will always be things we cannot prevent, and if we reflect too much on whether we should pursue new developments, then someone else will pick up the opportunity and do it in our place. As such, it seems that adopting a narrow focus to see tech innovation primarily as promoting productivity levels, also implies that ethical concerns—that might slow down progress—may need to be sacrificed.

All of these indicated that we are facing a double-edged sword where both slowing down the advancement of intelligent technologies versus not slowing down will make that the benefits of tech innovation are shared the least by those who need it the most. What to do? According to us, it is clear from history that progress cannot be stopped. If one slows down the pace of tech innovation in one nation or continent, then it will give an advantage to other nations, as such promoting power inequalities. Also, restricting tech innovation too much will ultimately not help anyone, and especially not the weaker groups of citizens where advances in technology can make a significant difference in their lives. For this reason, we propose an integrative approach where we start from the idea that we must continue to push forward tech innovation, but at the same time regulate its execution and install a stronger sense of moral awareness and responsibility among those in charge of the tech innovation journey. Below, we will summarize what specific actions may be needed to achieve this goal.

1. *Combine self- and government regulation*

   Advancement of AI has long thrived on a culture of lawlessness and irresponsibility as tech innovation needs to keep moving and who else besides the big tech companies are able to understand and monitor what is needed to maintain this level of progress [8]. The fact, however, that big companies, such as Facebook, Google, Microsoft, and Tencent, generate so much data while at the same time showing bad judgment with respect to and irresponsible management of those personal data (e.g., the Cambridge Analytica case), does indicate that we need to exercise more public control over these companies. For obvious reasons, in the past, the big tech companies have rarely advocated regulation of their inno-

vations, but their minds seem ripe today to engage in regulation by governments. Microsoft's President Bradford Smith, for example, asked the US congress directly to regulate the use of facial recognition technologies, which seems to indicate that the big tech companies realize that in their desire to marry maximizing profits and speeding up innovation efforts, ethical dangers exist that they themselves do not seem able or not willing to take care of.

Of course, Microsoft's request does hide a threat to the innovation process itself as well. If we restrict Western companies to stop using facial recognition systems, then an advantage will be given to, for example, China that sees no moral objections in using facial recognition systems to regulate people's finances, social credit scores and so forth. As such, we do note that we cannot overregulate to the point where it may inhibit innovation. In other words, we need to be careful that because of fear that ethical failures may emerge, we do not slow down tech innovation too much, but instead look for better solutions that can help prevent or manage those anticipated ethical failures.

To solve this dilemma, we argue that in addition to government regulations, we also need to leave room for self-regulation. Companies need to be given the opportunity to develop and implement a moral compass that can work in tandem with the regulation systems set out by the government. In other words, we need to create an ecosystem of tech innovation where one does not factor in ethical deliberations because one must but because one wants to. We do realize that proposing self-regulation as part of the solution to embrace tech innovation in morally appropriate ways can be considered naïve and even suspicious. After all, was it not the case that, for example, Facebook entertained for a long time that the market of digital platforms can regulate itself, thus not seeing the necessity for them to actively monitor how the platform was being used?

Mark Zuckerberg and others at Facebook indeed embraced Adam's Smith notion of the "invisible hand" to avoid responsibility to regulate the workings of the platform and instead attributed responsibility to the ones using the platform [9]. This was a perfect case of allowing for self-regulation, but instead of taking up the challenge to act as a moral leader who takes responsibility for the kind of consequences tech innovations may reveal, Facebook did not take an active stance and justified this by communicating that their only responsibility was to deliver the best technology possible. How end-users work with that technology was back then not seen as the company's responsibility. In the meantime, Facebook did have to learn the hard way that their perspective was not viable, but also completely irresponsi-

ble. As Zuckerberg noted on 4 April 2018 when he was speaking to journalists worldwide while the Cambridge Analytic case was unfolding: "We didn't take a broad enough view of what our responsibility was". The fact that Facebook failed to take this broader perspective at the same time underscores the failure of the founder himself, which he also acknowledged, by saying: "It was our mistake—it was my mistake." He further also noted that: "I started this place, I run it, I'm responsible for what happens here."

So, adding self-regulation as a force to our strategy to pursue tech innovation then also requires that companies commit to investments to develop the moral compass of their leadership and demonstrate that such ethical mindset is present and penetrates the work culture in its entirety. This brings us to our second point of action.

2. *We need more responsible leadership*

The business world is a volatile environment dominated by competitive tendencies. Within this ecosystem, time to reflect and deliberate difficult decisions for a longer time are usually considered a luxury. Time is money and both cannot be wasted. Companies deal with this corporate wisdom by only considering the "ethics" of a decision when the outcome is bad and represents a threat to their reputation. When outcomes are good, regardless of how it is achieved, responsibility to reflect on the ethics of the decision is less on their mind. It is a human tendency to focus primarily on outcomes, and only when the outcome falls short of expectations to evaluate the process that led to the undesirable outcome [10, 11]. This corporate reality makes that responsible leadership is usually demanded after things have gone wrong and only in rare occasions before things could go wrong. Anticipatory responsible leadership is given little leeway if no smoking gun can be provided demonstrating that ethical violations will happen for sure. In turn, if no smoking gun can be shown, it is difficult to use available financial budgets to scrutinize innovation because others will not recognize the need for a pro-active responsible leader. This kind of reasoning is illustrated nicely by the following thinking exercise [12].

Imagine that you are walking past a restaurant where you clearly see that the condition of the electric wiring in the kitchen is posing a serious threat to the safety of the people inside. It is clear to you that it does not take much for a fire to break out. Convinced about your assessment, you run into the restaurant and try to persuade people to leave the restaurant. You tell them that this is needed to protect their future health and survival. What will be the response? Most likely people will look at you in a bewildered way and think you have lost all your intellectual abilities. In other words, will they consider you to be a leader? Not really. It is more likely that

they will think you are a "zero" and not a "hero". Let us do this thinking exercise again but now imagine that, in a parallel universe, you are walking past the same restaurant again. But this time, a fire has now broken out in the kitchen and people eating there are under severe threat. Imagine you run in and save several people from the fire. What will be their response now? Most likely they will look at you as a leader. Yes, now you are not a "zero", but a "hero".

This thinking exercise makes clear that as humans we do not easily recognize the need for responsible leadership when nothing has gone wrong yet. However, when things do go wrong, we all want responsible leaders to rise to the occasion. This trend suggests that in their regulating efforts, governments and industries should pay more attention to incentivizing and evaluating targeted efforts for responsible leadership to be installed who are able to use an ethical point of view before actual decisions are made. We thus advocate that some part of regulations should include the obligation for companies to demonstrate a pro-active mindset that includes the requirement to consider ethical consequences before decisions are taken. Activities such as ethical risk sweeping need to become standard practices where leadership identifies ethical risks towards society and weighs these risks into their innovation strategies. Continuous education—as it is promoted nowadays as part of the new normal—as such will have to include a focus on developing leader's moral awareness and thinking that can help them to identify any moral potholes and act upon it ahead of making decisions.

3. *Bring AI faster to the people*

It is clear by now that the advancement of AI will not stop and continue pervading all dimensions of life. As a result, we cannot really slow down tech innovation. However, because AI could do so many great things for us, we do suggest that we should think more about how we can move AI advancement forward faster so that we all can enjoy its benefits. Although AI advancement is already taking place in a fast pace, it primarily serves the interests of the bigger companies, because tech innovation is mainly seen as the direct way to promote productivity and hence economic growth. However, many better, morally, and scientifically sound choices exist about what we can do with technologies, but unfortunately are not being envisioned yet. Because of our narrow focus on what optimization stands for, we are primarily focusing on those type of innovations that can create financial wealth quickly for a limited number of stakeholders, while missing out on innovations that can enrich our cultural, artistic, and social needs in better ways. As a matter of fact, we think that AI can also make us more human by augmenting our soft skills and unique human

experiences in more diverse ways [13]. To make this kind of enrichment happen, we need more AI applications in diverse areas of society and push forward the advancement of intelligent technologies to create value for humanity across several dimensions of life. Again, to make AI happen faster for all people, we will need leadership and governance that adopts a broader framework than only optimizing economic growth. In essence, today—more than ever—we need leadership asking the right kind of questions that can help improve our lives beyond financial interests.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. De Cremer, D.: Leadership by algorithm: who leads and who follows in the AI era? Harriman House, London (2020)
2. Pwc.: PwC's global Artificial Intelligence study: Exploiting the AI revolution. (2018) Retrieved from: https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html
3. PwC.: 22nd Annual Global CEO Survey. (2019) Retrieved from: https://www.pwc.com/gx/en/ceo-survey/2019/report/pwc-22nd-annual-global-ceo-survey.pdf
4. Brynjolfsson, E., McAfee, A.: The second machine age: work, progress, and prosperity in a time of brilliant technologies. W.W. Norton & Company, New York (2016)
5. Keynes, J.M.: A treatise on money. Harcourt. Brance and Company, New York (1930)
6. Spocchia, G.: Jeff Bezos criticised by Amazon workers and customers after thanking them for funding space launch. (2021). Retrieved from: https://www.independent.co.uk/news/world/americas/amazon-workers-slam-jeff-bezos-b1887944.html
7. Vallor, S.: Mobilising the intellectual resources of the arts and humanities. (2021). Retrieved from: https://www.adalovelaceinstitute.org/blog/mobilising-intellectual-resources-arts-humanities/
8. Nemitz, P.: Constitutional democracy and technology in the age of artificial intelligence. Philos Trans R Soc A: Math Phys Eng Sci (2018). https://doi.org/10.1098/rsta.2018.0089
9. De Cremer, D.: Why Mark Zuckerberg's leadership failure was a predictable surprise. The European Business Review, May–June, 7–10 (2018)
10. Baron, J., Hershey, J.C.: Outcome bias in decision evaluation. J. Pers. Soc. Psychol. **54**, 569–579 (1988)
11. Allison, S.T., Mackie, D.M., Messick, D.M.: Outcome bias in social perception: Implications for dispositional inference, attitude change, stereotyping, and social behavior. In: Zanna, M.P. (ed.) Advances in experimental social psychology, vol. 28, pp. 53–93. Academic Press, New York (1996)
12. De Cremer, D.: What are we doing today to prevent our company's next ethical disaster. The European Business Review, January–February, 7–10 (2020b)
13. De Cremer, D., Kasparov, G.: The ethical AI—paradox: why better technology needs more and not less human responsibility. AI Ethics (2021). https://doi.org/10.1007/s43681-021-00075-y