
Rise of the (Good) Machines: A Blueprint for Action

Title: Rise of the (Good) Machines: A Blueprint for Action

Author: James Stauch and Alina Turner

Published in: *The Philanthropist*

ISSN: 2562-1491

Date: March 16, 2020

Original Link: <https://thephilanthropist.ca/2020/03/rise-of-the-good-machines-a-blueprint-for-action/>

Date of PDF Download: March 29, 2020

This is the second of [two articles](#) about artificial intelligence.

"An AI utopia is a place where people have income guaranteed because their machines are working for them. Instead, they focus on activities that they want to do, that are personally meaningful like art or, where human creativity still shines, in science."

- Oren Etzioni, CEO, Allen Institute for Artificial Intelligence

While some argue that new jobs will always emerge in place of old ones – an axiom that has, so far, been borne out – we are witnessing an acceleration of technology-induced employment displacement. Flesh and bone versions of bank tellers, customer service reps, stock and bond traders, and telemarketers are already becoming scarcer. One study predicts that artificial intelligence (AI) will outperform (and soon thereafter replace) translators by 2024, truck drivers by 2027, retail sales workers by 2031, fiction authors by 2049, and surgeons by 2053.^[1] Even those involved in employment-intensive processes such as oil sands production are trumpeting the risk-mitigation and shareholder benefits of “de-manning” operations.

In our previous article, we spoke of AI in the context of the Fourth Industrial Revolution. But AI may well be the most important disruption to human social organization since the Agricultural Revolution. And yet, the very sector charged with monitoring, advocating for, and improving social impact and civil discourse, is barely talking about it. This essay is an appeal to those working at non-profits and philanthropic organizations to start paying much closer attention to AI.

News about AI tends to focus on its shadow side: the built-in blind spots and cognitive biases that amplify into serious racial, gender, and other forms of discrimination. Or the commodification of our personal information, rolled up into big data analytics and the forensic manipulation of social media platforms to spread misleading memes, fake news, and other forms of poisonous and counterfeit discourse. However, as mathematician Cathy O’Neil points out, the algorithms are not responsible for the societally-destructive effects of machine learning – it’s the human bias implanted in algorithm coding.[\[ii\]](#)

There is another side to AI, one filled with the potential to enhance and accelerate our desire for a more caring, creative, just, and sustainable world, as well as – counterintuitively – the flourishing of human potential. As shorthand, we’ll call these many and diverse aims “social good.” And there is a crescendo of activity underway – in R&D, pilot testing, and nascent applications – in the domains of health, the environment, arts, education, and other social good activities.

To date, we have seen a trickle of news stories, expert reports, and opinion about AI and its role in social good, but this will soon turn into a torrent. The first major report on AI for social good, produced by the McKinsey Global Institute, came out in 2018.[\[iii\]](#) Since then, there have been major international gatherings on this topic in Geneva (sponsored by XPRIZE), Paris (at the invitation of French President Emmanuel Macron), and New York (on using AI to advance the UN Sustainable Development Goals).[\[iv\]](#) But what is happening on the ground and how might it affect our everyday work?

Health care is currently where we see some of the most profoundly disruptive AI applications, as well as applications that are approaching widespread use. In diagnostics in particular, neural network technology has matched or exceeded the accuracy of specialists in detecting skin cancer, pediatric diseases, and more than 50 different eye disorders. Radiology, pathology, ophthalmology, and cardiology are just some of the fields using deep learning algorithms in diagnostics. AI innovator Neil Jacobstein predicts that “we will soon see an inflection point where doctors will feel it’s a risk to not use machine learning and AI in their everyday practices because they don’t want to be called out for missing an important diagnostic signal.”[\[v\]](#) AI will help practitioners analyze and synthesize complex data regarding symptoms, risk factors, and drug interactions into a more patient-tailored treatment regimen. Treatments will be less invasive and more bespoke; many assistive devices will be exponentially more effective; health care administration will be streamlined; and prognoses for a wide range of diseases will improve with AI-assisted medicine. Patients will also be more empowered with AI-assisted phone apps: we already have an array of apps, for example those that allow you to scan your own skin with to detect melanoma.

Environmental protection is another realm in which AI could positively transform the status quo. Machine learning and [deep neural networks](#) will be able to help societies and governments enhance predictive modelling for climate change, both around the severity and precision of expected impacts (for adaptation purposes), but also for prioritizing mitigation strategies. AI is already being used to develop clean tech such as solar panels, batteries, and materials that can absorb emissions or conduct photosynthesis. The McKinsey study chronicles 160 AI applications already in use or in development, across all 17 UN Sustainable Development Goals.[\[vi\]](#) When we factor in another tech development – quantum computing – we are not far off from having artificial general intelligence process more physics papers in one second than a human could complete in a thousand lifetimes. When one considers this potential, the prospect

for new and cheaper sources of energy, new modes of transportation, or new methods of carbon sequestration, for example, will be within closer reach. Beyond climate change, there are many other environmental applications of AI currently in beta testing or early deployment, often paired with other technologies, such as drones to curb poaching, blockchain to catalogue species, or sensors to conserve soil and water in food production. The Energy Futures Lab, convened by The Natural Step Canada, has developed the Energy.AI program to help accelerate collaborative initiatives that apply machine learning to low-carbon energy production.

Meanwhile, in the **arts**, AI has already been used to compose music, generate images, write poetry and film scripts, and even dance. The team at the Expressive Machinery Lab at Georgia Tech have developed a method called "collaborative movement improvisation," learning from human dance partners. We are also seeing the emergence of the "creative technologist," artists who utilize AI to augment or enhance their work. One such creative technologist, for example, developed a machine learning homage to Jack Kerouac, in the form of an AI-generated novel called *1 the Road*. Taishi Fukuyama, the creator of Amadeus Code, reminds us that major technological advances have always ushered in exciting new artistic frontiers.^[vii] Technological advances also tend to democratize, decentralize, and dematerialize production, as we can expect with the widespread use of "generative adversarial networks," the technology responsible for generating "deep fakes" (simulated photographs and video). The arts are also critical for AI development to be more humane and, well, human. Augur, a Stanford University project, interprets a large body of literary fiction to better understand everyday human behaviour and emotional responses. Again, AI's power is amplified in combination with other technologies like augmented, virtual, and mixed reality, which are already being used in retail, e-commerce, and architecture, but also increasingly in photography, cinematography, real estate, driving, camping, gardening, and in countless other pursuits where the imagination of artists and creative professionals will be in very high demand. As Tad Friend notes in *The New Yorker*, our future "will need experts in unexpected disciplines such as human conversation, dialogue, humor, poetry, and empathy."^[viii] There has never been a better time to major in drama or philosophy, or a worse time to major in accounting or finance (with the exception, by the way, of the still-infant sub-field of social finance).

Within **education and research**, the benefits of AI will almost certainly outweigh the downsides of disruption. Northeastern University President Joseph Aoun argues in his book *Robot-Proof: Higher Education in the Age of Artificial Intelligence*^[ix] that the role of education must not simply be to understand how we can use better data and machine learning, but also to fill needs in society that even the most advanced AI agent cannot. For those who champion the liberal arts, or "general education," we may be on the cusp of a new golden age insofar as the liberal arts help interpret, navigate, and gain agency amid growing complexity, diversity, and change. AI is already disrupting the factory model of education through personalized, adaptive applications that tailor learning to students' specific needs, passions, and learning styles. Predictive AI has already saved student lives in pilot projects that monitor student texting activity or in-school internet activity, flagging and triangulating complex risk factors for suicidal behaviour, leading to more timely intervention. Within research, we have seen many examples of AI-generated academic papers (the first auto-generated wave of papers mainly shone a light on the scourge of predatory journals or academic conferences that have very low submission standards). The second wave of AI-generated academic writing is far more serious and disruptive. Consider, for example, the Applied Computational Linguistics Lab (ACoLi) at Goethe University in Frankfurt, which published an AI-authored textbook on the subject of lithium ion batteries, distilling insights from more than 53,000 peer-reviewed papers published within the

last three years into a 180-page meta-analysis.[\[x\]](#)

There are countless other social good applications possible, or already in development. We are seeing non-profit organizations in the US analyze open source data to report and rate racially-motivated police harassment. Big data analytics and machine learning are being used to identify food deserts, underserved communities with respect to educational or transit access, high-cost users of public services, or families at greater risk of homelessness. One experiment that looks at predictive factors for riots, lynchings, and other mob violence in Liberia is producing useful and surprising insights thanks to machine learning techniques.[\[xi\]](#) Meanwhile, AI deep learning applied to crisis counselling conversations has enabled a reshuffling of queues and adjustment of counselling techniques to help save more lives.

In social innovation parlance, we speak of “complex adaptive systems” to characterize so much of the work in the non-profit milieu, especially with respect to community services and environmental protection. This complexity has historically meant that technology was of marginal utility, relative to public policy work or nuanced human relationships and interventions. But AI advances are changing this equation. Complex algorithms on suicide prevention, for example – for which there are hundreds of risk factors in the mix, re-calibrated to specific cultural and geographic contexts – are showing huge promise in beta testing.[\[xii\]](#)

Despite all these many exciting advances, biases and blind spots remain in abundance. Moreover, the vast majority of AI development is happening in realms that are not focused on social good, but instead for commercial purposes or (particularly in China) for the purposes of state surveillance and control. We are at an important crossroads, and we have little time to choose a path that will enhance, not degrade, the social wellbeing, democratic vitality, and natural ecosystems on which we rely. It is imperative that we take some critical steps in Canada to help us overcome fears of AI and gain agency and voice in how it is developed:

1. *Inclusive research and development:* Canada is a leader in both social innovation and tech innovation, and it is time for these two worlds to collide and connect. It is not enough for the tech community to assume they have a handle on social good – this is how biases and blind spots emerge. The development of machine learning systems and AI should not only include computer scientists, software engineers, and data scientists, but also sociologists, philosophers, anthropologists, human rights lawyers, economists, historians, and many other vocations in an attempt to better understand the long-term effects and potential of AI on social good. Moreover, the end users or clients of social interventions ought to be at the table too, not merely experts and service providers. If there is one area where human-centred design must be a minimum specification, it is in AI development. Mark Latonero, a Technology and Human Rights Fellow at Harvard’s Carr Center, presciently cautions that “the fanfare around these [AI] projects smacks of tech solutionism, which can mask root causes and the risks of experimenting with AI on vulnerable people without appropriate safeguards . . . Tech companies that set out to develop a tool for the common good, not only their self-interest, soon face a dilemma: They lack the expertise in the intractable social and humanitarian issues facing much of the world.” Noting some of the new partnerships that have emerged between tech companies such as IBM and Google, and non-profits like National Geographic Society and the Leonardo DiCaprio Foundation, Latonero concludes that “partnerships are smart. The last thing society needs is for engineers in enclaves like Silicon Valley to deploy AI tools for global problems they know little

about.”[\[xiii\]](#) We are encouraged by recent announcements like CIFAR’s [AI and Society program](#) and the newly-announced RBC Data Analytics and Artificial Intelligence Project at Western University, which will focus on “answering big questions for the good of society.” The Schwartz Reisman Innovation Centre at the University of Toronto promises to look into the intersection of socially-responsible biomedicine and AI. But so much more needs to be in place, inside and outside academia, from coast to coast to coast. We reference many other promising practices at universities around the globe in *In Search of the Altruism*, but one in particular is worth highlighting: University of Southern California’s Center for Artificial Intelligence in Society, which is run by a social work professor and employs healthcare practitioners, policy analysts, and social sector thinkers and doers alongside tech engineers and developers.

2. *Disruption tolerance*: As previous industrial revolutions have shown, the social sector is far from disruption-proof. For example, membership fee-dependent knowledge networks are generally not viable in a world of free and ubiquitous information. Co-creation of an AI-accompanied future will require social sector leaders to have a stronger voice in public decision-making, and embrace risk and deeper forms of system-wide collaboration. Many commentators have urged social practitioners to carry the torch for evidence-based, data-driven solutions, but are we prepared for where this may take us? AI may well prove to be the worst nightmare of status quo politicians – or status quo non-profits and funders – as we consider the implications of social intervention optimization informed by deep learning. As we muse in our report *In Search of the Altruism*, “Superintelligent AI may well determine, based on reams of high-quality peer-reviewed research and petabytes of liberated data on pilot projects and social intervention prototypes, that we need policies and programs that are politically unpalatable in today’s context. Universal basic income, a flexible 15-hour workweek, decriminalization of all narcotics, psychotropic treatment of addictions, nature-based incarceration, or any number of other audacious-sounding social good decisions may emerge.”[\[xiv\]](#)
3. *Digital commons*: Capitalism itself will be profoundly disrupted for social good outcomes to flourish under an AI-dominated future. In McKinsey’s analysis of barriers to AI for social good, data inaccessibility is the most pressing issue. Philanthropy futurist Lucy Bernholz adds that “if we want to keep measuring civil society activity – giving, volunteering, activism, participation, etc. – we need to make sure the data on our collective actions are not locked down by proprietary platforms.”[\[xv\]](#) To this end, Canada’s [Digital Charter](#), announced in May of 2019, and the federal government’s [Open Resource Exchange](#), are welcome developments in society’s necessary battle for an AI commons. While smaller- and medium-sized charities and non-profits experienced a digital divide vis-à-vis the private sector through the 1990s and early 2000s, there was a parallel renaissance underway in the form of open source coding (e.g. Linux), Peer-to-peer file sharing, Creative Commons, and wikis, which are all manifestations of a digital commons. Funders must support such open and collaborative initiatives – Open North and Data for Good being two such examples – as well as watchdogs and data activists.
4. *Hyper-citizenship*: Canada is fortunate to have many organizations working to enhance our democratic engagement and institutions. Samara, CIVIX and Democracy Watch are just a few examples. But we need a new skillset on top of basic civic literacy – call it extreme enlightenment or hyper-citizenship. This involves distinguishing counterfeit

news, knowledge, and movements from real (i.e. authentic and rigorous) news, knowledge, and movements. The rapidly advancing sophistication of deep fakes and artificial content is, so far, outpacing AI's ability to detect it. At the same time, we are starting to think more critically about the role of "empathy" in care-based decision-making. As we assert in *In Search of the Altruism*, drawing inspiration from Paul Bloom's notion of "rational compassion," "Empathy biases the near and familiar over the different and far-away. It is a useful and necessary mental function, essential to our very sense of humanness, but it is also the same region of the brain that produces racism, parochialism and wildly uneven — often deeply irrational — social outcomes when extended to the practice of charity or public policy."

5. *Creativity and collective imagination*: Marshall McLuhan said that art is a "distant early warning system." In a similar vein, in a recent Walrus Talk, Canadian futurist Hamoon Ekhtiari said we need to go beyond building an ethical framework and social purpose orientation for AI, urging us to foster a collective imagination in order for AI to truly serve humanity.[\[xvi\]](#) Creative mindsets, systems thinking, ethical inquiry, and mental elasticity will all be needed in abundance, suggesting, once again, that the arts and humanities may be more critical than ever. How might we build effectively altruistic algorithms premised on global fairness around relief of poverty? Or code in the "veil of ignorance" to the ovarian lottery, as Warren Buffet has called it, removing barriers to social mobility? Or feminize AI decision-making?
6. *Public commitments and protocols*: As we assert in *In Search of the Altruism* "If the monopoly power of commercial tech giants and totalitarian regimes is not greatly circumscribed, AI will fail to serve the common good."[\[xvii\]](#) Canada is home to one of the most important public commitments to socially responsible AI, the [Montreal Declaration](#). The Swedish Future of Life Institute has catalyzed a similar commitment across the Baltic states, and Canada has been at the forefront of a number of critical processes, such as the [International Panel on Artificial Intelligence](#) (IPAI), which is to include representation from civil society and will align AI investment to the UN Sustainable Development Goals. CIFAR and the Brookfield Institute have hosted a series of public policy conversations on AI, which have included non-profit sector leaders, among others.
7. *Tech and data literacy*: Notice how far down the list this is, when it might seem like the most obvious and urgent. Certainly, we believe that social changemakers should generally be much more curious about AI and technological advancements. Non-profit managers, designers, and evaluators should consider enrolling in AI bootcamps, online courses and short intensive development opportunities for non-tech professionals. But we don't all need to code. We simply need to be basically literate and have generalized familiarity. Most importantly, we should not approach technology with fear and reflexive Ludditism. We must take steps to source, hire, and attract tech talent, which in turn means that non-profit work must be less precarious and more dignified and properly remunerated. Our cheap and cheerful approach to non-profit resourcing also means that new technologies, software, and IT is too-often second tier or second hand. At the same time, universities should create social impact work-integrated residencies for AI-focused grad students. Tech companies need to reciprocally up their inclusion game: for example, only 10% of Google's employees working on "machine intelligence" are female. We can take inspiration from the Canadian STEM non-profit organization Actua,

which is bringing AI content into high schools, as well as from initiatives south of the border such as the Partnership on AI and AI4ALL, which aim to make AI approachable for the general public and underrepresented groups. Diverse, interdisciplinary, cross-functional teams that bridge the social/tech divide are critical, as are "data translator" NGOs that employ data scientists who can interpret and stress-test an algorithm's "brittleness" and bias vulnerability. As we optimize AI for given social objectives, that objective must incorporate everything we care about, and everything that is important to the citizens, creatures, or ecosystems being served or "helped."

Right now, the worlds of social innovation and tech innovation are speaking completely different languages and largely talking past each other. This must end. As Bernholz brilliantly observes, "There is no 'clean room' for social innovation — it takes place in the inequitable, unfair, discriminatory world of real people. No algorithm, machine learning application, or policy innovation on its own will counter that system and it's past time to keep pretending they will." [xviii] We need mediated and new "collaborative spaces where the relationships, trust and shared platforms for ideas and experiments can flourish, and where radical, even revolutionary, action can emerge." [xix] We need many more shared conversations, drawing inspiration from the Civil Society Futures project in the UK, as well as scenario planning, strategic foresight and government-supported AI and common good research.

The vigilant, cautious, and creative amalgamation of machine super-intelligence with human learning has enormous potential to help us solve "wicked problems." But is civil society sufficiently future-proof (or future-ready), when the stakes are so stratospheric, and the rewards and risks challenge the very bounds of our imagination? Nick Bostrom, director of the University of Oxford's Future of Humanity Institute, concludes – perhaps hyperbolically – that "Machine intelligence is the last invention that humanity will ever need to make. Millennia from now, people will look back and note that the one thing we did that really mattered, we got right." He refers elsewhere to the notion of a "detonation" as being a more accurate metaphor. Getting this invention right, or setting the detonation charges in the optimal way, may well be the most vital social and philanthropic goal we can pursue. Earth is entering uncharted waters and, as the machines rise, we have never had a greater need for all hands on deck.

[i] Grace, K., Salvatier, J., Dafoe, A. Zhang, B. and Evans, O. (2018). When Will AI Exceed Human Performance? Evidence from AI Experts, *Journal of Artificial Intelligence Research*.

[ii] O'Neil, C. (2018) Algorithms are not truth machines (animation). RSA. Retrieved from <https://www.youtube.com/watch?v=heQzqX35c9A>

[iii] Chui, M., et al. (2018, December) *Notes from the AI Frontier: Applying artificial intelligence for social good* (discussion paper). McKinsey Global Institute.

[iv] Latonero, M. (2019, November 18) AI for Good is Often Bad, in *Wired*. Retrieved from <https://www.wired.com/story/opinion-ai-for-good-is-often-bad/>

[v] As quoted in Diamandis, P. (2019) How Artificial Intelligence Will Reinvent Industries, Products and Services (blog post) *21st Century Tech*. Retrieved from <https://www.21stcentech.com/artificial-intelligence-reinvent-industries-products-services/>

[vi] Chui, et al. (2018).

[vii] Hu, C. (2018, December 10). The Future of Entertainment: Can Robots Build Better Hits?, in Rolling Stone. Retrieved from <https://www.rollingstone.com/culture/culture-lists/future-entertainment-technology-music-tvmovies-760659/ai-songwriting-760685/>

[viii] Friend, T. (2018, May 14). How Frightened Should We Be of A.I.? in *The New Yorker*.

[ix] Aoun, J. (2018) Robot-Proof: Higher Education in the Age of Artificial Intelligence. Boston: MIT Press.

[x] Beta Writer (2019). Lithium-Ion Batteries: A Machine-Generated Summary of Current Research. Springer Nature

Switzerland.

[xi] Blair, R.A., Blattman, C. and Hartman, A. (2017, March). Predicting Local Violence: Evidence from a Panel Survey in Liberia. *Journal of Peace Research*, 54, no. 2: 298–312.

[xii] Anderssen, E. (2018, Nov. 24) Can an algorithm stop suicides by spotting the signs of deep despair? *The Globe and Mail*. Retrieved from <https://www.theglobeandmail.com/canada/article-can-an-algorithm-stop-suicides-by-spotting-the-signs-of-despair/>

[xiii] Latonero (2019).

[xiv] Stauch, J. Turner, A. and Escamilla, C. (2019) *In Search of the Altruism: AI and the Future of Social Good*. Institute for Community Prosperity, Mount Royal University.

[xv] Bernholz, L. (2019). *Philanthropy and Digital Society: 2019 Blueprint*. Stanford Center on Philanthropy and Civil Society, 2019. Retrieved from <https://pacscenter.stanford.edu/publication/philanthropy-and-digital-civil-society-blueprint-2019/>

[xvi] Ekhtiari, H. (2018, Oct. 3) Audacious Futures. *Walrus Talks Humanity and Technology* (Toronto). Retrieved from <https://www.youtube.com/watch?v=fhPv8uNlxRs>

[xvii] Stauch, Turner and Escamilla (2019).

[xviii] Bernholz, L. (2018) Flipping our algorithmic assumptions (blog post). Stanford Center on Philanthropy and Civil Society. Retrieved from <https://medium.com/the-digital-civil-society-lab/flipping-our-algorithmic-assumptions-98baef2d5cbd>

[xix] Stauch, Turner and Escamilla (2019).