IN SEARCH OF THE ALTRUISTHM: AI and the Future of Social Good

James Stauch, Alina Turner and Camilo Camacho Escamilla
<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>How did we get here?</td>
</tr>
<tr>
<td>5</td>
<td>Why should we care?</td>
</tr>
<tr>
<td>7</td>
<td>Imagining the future: Where do we fit in?</td>
</tr>
<tr>
<td>9</td>
<td>A closer look at AI's social good applications &amp; implications</td>
</tr>
<tr>
<td></td>
<td>AI and health</td>
</tr>
<tr>
<td>11</td>
<td>AI and Environmental Protection</td>
</tr>
<tr>
<td>12</td>
<td>AI, Arts &amp; Creativity</td>
</tr>
<tr>
<td>13</td>
<td>AI and Education</td>
</tr>
<tr>
<td>14</td>
<td>AI and other Social Good Applications</td>
</tr>
<tr>
<td>16</td>
<td>AI and Democracy</td>
</tr>
<tr>
<td>17</td>
<td>AI in the service of social good: How do we prepare?</td>
</tr>
<tr>
<td></td>
<td>1. An AI commons</td>
</tr>
<tr>
<td>18</td>
<td>2. Hyper-citizenship</td>
</tr>
<tr>
<td>19</td>
<td>3. Creativity and collective imagination</td>
</tr>
<tr>
<td>20</td>
<td>4. Technological and data literacy</td>
</tr>
<tr>
<td>21</td>
<td>5. Disruption tolerance</td>
</tr>
<tr>
<td>22</td>
<td>6. Public commitments, declarations and protocols</td>
</tr>
<tr>
<td>23</td>
<td>An opportunity for world-leading integration of social innovation and AI development</td>
</tr>
<tr>
<td>24</td>
<td>AI and the Future of Good</td>
</tr>
<tr>
<td>25</td>
<td>About the Authors</td>
</tr>
<tr>
<td>26</td>
<td>About the Institute</td>
</tr>
<tr>
<td></td>
<td>Acknowledgements</td>
</tr>
<tr>
<td>27</td>
<td>References</td>
</tr>
</tbody>
</table>
"With an anxiety that almost amounted to agony, I collected the instruments of life around me, that I might infuse a spark of being into the lifeless thing that lay at my feet."

- Mary Shelley, Frankenstein
Humanity is in the midst of a crucial social and economic transformation. As World Economic Forum Founder Klaus Schwab (2016) frames it, whether we know it or not, we have entered the early stages of the Fourth Industrial Revolution. In much the same way as steam engines, internal combustion, and the internet were among the innovations that marked the First, Second, and Third Industrial Revolutions respectively, this Fourth Revolution is marked by machine-learning, the Internet of Things (IoT), autonomous vehicles, 3-D printing, blockchain, Big Data, gene editing, implantable devices and, potentially, quantum computing. These technologies blur the lines between the physical, digital, and biological realms, as well as between human and machine cognition (Schwab, 2016).

The leitmotif that pervades most of these ever-shifting technologies is the rise of Artificial Intelligence, or AI. The concept of AI was coined in 1956 after computer scientist John McCarthy convened a conference on AI, which he described as "thinking machines." At the time it was only theoretically conceivable insofar as "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" (Moor, 2006).

We are still in the early, infant steps of ‘weak’ or ‘narrow’ AI and many years, or decades, away from realizing the Holy Grail of Artificial General Intelligence, or AGI. This is when DeepMind co-founder and world-leading AI developer Demis Hassabis asserts that we can expect AI to embark on rapid, unsupervised learning, and, in the process, master a variety of disciplines such as "cancer, climate change, energy, genomics, macro-economics [and] financial systems" (Hodson, 2019).

Although the consequences of a world dominated by AI are difficult to predict, the rise of AI does not have to be a terrifying prospect. As this paper will argue, AI can be made to be generative, beautiful, and not merely ethical, but rationally compassionate and just, enriching our lives beyond what we can currently imagine. But it will only do so if civil society — including citizens living at the margins, and people involved in social good pursuits, from artists and teachers to health professionals and social workers — become much more interested and involved.

This is not some exotic, far off prospect. AI is already here in many forms. But our future is still in our hands, at least for the next decade or two. Nick Bostrom, Director of the Future of Humanity Institute at Oxford University, notes that "Machine intelligence is the last invention that humanity will ever need to make” (Bostrom, 2015). Getting this invention right, from a social responsibility standpoint, may be the most important existential, public policy and social goal we can possibly pursue.

1A precise definition of AI is elusive, in part because there is no agreed-upon definition of "intelligence".
It is important to remember that as futuristic and sci-fi as some of the public discourse on AI might seem, the building blocks of Artificial Intelligence have been with us for about a century already. The early advancements made by Alan Turing in the 1930s enabled computers to contain large amounts of information and machines to simulate many intellectual tasks. In the 1950s, Machine Learning emerged as a subfield of computer science, focused on teaching machines how to respond to certain tasks through algorithms and pattern recognition. At this early stage, machines were not only able to “think” by themselves, within a narrowly confined domain, but also to learn how to perform simple (from a human perspective) intellectual tasks without being specifically programmed to.

Fast-forward to the 1990s, which saw the beginning of the broadly public version of the internet (via the World Wide Web and e-mail), the ubiquitous use of hard disk storage, the creation of the Linux kernel (which sparked a wave of open source innovation and peer to peer sharing), and the boom of data centres amid the dot-com bubble. In 1997, IBM’s Deep Blue famously defeated World Champion Gary Kasparov in chess. Rapid growth in the efficiency of computing enabled the storage and transfer of vast amounts of data, appropriately dubbed Big Data (Jordan & Mitchell, 2015). If algorithms are the DNA of AI, data is its food and oxygen. In 2019, the size of the global ‘datasphere’ reached 40 zettabytes (40 trillion gigabytes) and is rising exponentially.

The rise of deep reinforcement learning and artificial neural networks, embodied in innovations such as AlphaGo Zero, now self-teach and ultimately perform superhuman tasks in challenging domains, essentially free of human input (Asghar, 2016). This new era of ‘meta-learning’, where machines themselves learn how to learn, is key in presaging a near-future AGI. The last five years have also witnessed a convergence between brain science and computer science. AI capabilities such as natural language processing (including text-to-voice, language translation and comprehension), object and facial detection and recognition, are already shrinking the human-machine cognition gap.

The speed and intensity of these innovations is further amplified by rapidly growing investment in AI development. Globally, investment in AI-focused companies jumped from $589 million in 2012 to $5 billion in 2016 (Brynjolfsson, Rock, & Syverson, 2017). According to PWC, AI will add $16 trillion (US) to the global GDP by 2030 (Street, 2019).

From a technological innovation standpoint, Canada is very much in the AI game. The federal government announced $125 million investment in a Pan-Canadian Artificial Intelligence Strategy to retain and attract top academic talent, increase the number of postgraduate trainees and researchers studying AI and deep learning across Canada’s main centers of expertise in Montreal, Toronto-Waterloo and Edmonton (GOC, 2017). Among the world’s leading AI scientists are Turing Award recipients Geoffrey Hinton at the University of Toronto and Yoshua Bengio of the Université de Montreal. The University of Alberta now ranks third in the world for artificial intelligence research, behind only Carnegie Mellon University and Tsinghua University in China (Staples, 2018). However, Canada is not on the leading edge with respect to AI and the future of social good. This might be starting to change, as in recent months we have seen the creation of the International Observatory on the Societal Impacts of Artificial Intelligence and Digital Technologies at the University of Laval, the University of Guelph’s Centre for Responsible and Ethical AI, and the University of Toronto’s announced Reisman Institute for Technology and Society (Venne, 2019). This paper offers a framework and some ideas for what else we might do ensure AI works for social good in Canada.

\(^2\)AlphaGo was the first computer application to defeat a world champion at the Chinese game of Go. AlphaGo Zero is the fourth iteration of this program, and by far the most sophisticated, also using a fraction of the power consumption of its early predecessors.
At some point in the coming decades, we are expected to reach the "technological singularity," the point at which AI abruptly triggers runaway technological growth, resulting in unfathomable changes to human civilization (Edan and Moore, 2012). The divergent path scenarios such a singularity could take humanity range from the harrowing enslavement or extinction of our species to the exhilarating possibility of becoming an incalculably more powerful and omniscient species, perhaps even conquering mortality itself.

It would be naïve, if not downright ignorant, to think that the social sector is not already profoundly impacted by these changes. In certain domains, the rate, scope, and depth of machine learning appears to be outpacing the expansion of human insight and awareness. Our daily lives are algorithmically-augmented and assisted in countless ways already, from our Netflix-viewing or Spotify-listening preferences, to predictive search results, news feeds and chat bots, to ride sharing and air transportation (Markoff, 2015). Since its introduction in 2014, over 100 million Alexa personal assistant devices have been sold (Matney, 2019). We have already seen the practice or potential for racial and gender bias revealed in algorithms for hiring, policing, judicial sentencing, and financial services (Villasenor, 2019; Buranyi, 2017). We have even borne witness to socially malicious applications of AI, such as the use of bots to generate and amplify anti-vaxxer social media posts (Subrahmanian et al., 2016). Beyond this, AI is already provoking disruption to employment patterns and the job market, likely intensifying patterns of inequality, at least in the near term (Ernst, Merola and Samaan, 2018).

From a social impact standpoint, certain aspects of AI are inarguably beneficial. For example, Facebook’s ‘proactive detection’ technology identifies self-harming and suicide-risk behaviours (Constine, n.d.). Machine Learning is being applied to make the lives of those impacted by disabilities much easier: Voice recognition and speech-to-text come to mind. Instant translation is another example that is breaking down communication barriers across languages with positive social impacts³. But, as commentators like Amy Webb (2019), Franklin Foer (2018) and Zeynep Tufekci (CBC News, 2018) point out, AI’s application to date — including Facebook’s very business model — typically has a profit imperative with consequences that are rightfully classified as dystopian in their logical implications. As Webb observes, “The future of AI — and by extension, the future of humanity — is already controlled by just nine big tech titans, who are developing the frameworks, chipsets, and networks, funding the majority of research, earning the lion’s share of patents, and in the process mining our data in ways that aren’t transparent or observable to us” (Webb, 2019). And many supposed ‘upsides’ of AI, from a utilitarian perspective, have unsettling downsides, from a human rights and ethics perspective. Take, for example, the use of an AI called DeepGestalt (Gurovich, et al, 2019), which uses facial recognition to detect genetic disorders among the general population. Clearly beneficial from a health care standpoint, but with troubling ethical implications.

Philanthropy futurist Lucy Bernholtz cautions that social impact organizations, including foundations and nonprofit service providers, need to become acutely aware of these phenomena:

> The social media systems are purpose-built to manipulate. Facts and good intentions aren’t enough. Understanding the nature of the information ecosystem — the ways it makes getting your message heard harder, not easier, and the ways it threatens the well-being and safety of those you are trying to help — is no longer an optional, edge requirement. It’s reality for all of us in the digital age (Bernholz, 2017).

The emergence of the social sector as we know it over the last two centuries is a direct result of the recognition that technological advancements are not universally beneficial. The development of new technologies has brought important advancements for human prosperity and health, but have also impacted quality of life negatively for many. The three previous industrial revolutions were

massively disruptive, giving rise to significant social challenges and market externalities such as human-induced climate change (which began when we started burning coal en masse), rapid urbanization, child labour, urban sanitation issues, air and water pollution and industrial-scale farming and incarceration. The first industrial revolution gave rise to social reformers like Robert Owen, Elizabeth Fry and John Howard, revisions to the Poor Laws in England, the genesis of the modern charitable sector, urban planning and corporate social responsibility, among other double-edged innovations.

Despite the increasing productivity and economic growth during the Second Industrial Revolution, social and economic disparities intensified. In response, we saw the emergence of formal charitable organizations and peer-support efforts such as the YMCA, Salvation Army, and the Red Cross. The Pemsel Case, the 1891 decision of the British House of Lord, outlined the ‘four pillars’ of charity still in use today. In Canada, the 1914 establishment of bodies such as the Neighborhood Workers Association aimed to coordinate the work of such efforts (Elson, 2009, p. 41). The same decade saw the emergence of the Antigonish Movement in Nova Scotia, blending adult education, co-operatives, microfinance and community development (Lotz, 2005). The emergence of such organizations — the start of what we now call the “social economy” — profoundly changed how society thought about and dealt with social good.

By the Third Industrial Revolution, these organizations were recognized by the state as essential to alleviating the effects of government/market failures. They were joined and fueled in part by the rise of mass media technologies in the 1960s and 1970s by civil rights groups, ethno-cultural societies and Indigenous and environmental organizations. In Canada, this “Third Sector” ballooned: between 1974 and 1990 the number of registered charities grew by 80% from 35,113 to 63,186 (Elson, n.d. p. 13). While many areas of social need continue to grow, the combined value of government, charitable and corporate support for the innovation and development of the social sector has not kept pace. As a result, we now see a structural social deficit (Emmett, 2017), with intensified calls for investments in social innovation, social finance, social R&D, adaptive capacity and systems leadership.

Just as the very existence of the social sector is a byproduct of past industrial revolutions, we should expect the Fourth Industrial Revolution to have tremendous implications. Already, we have seen AI displace entire professions including bank tellers, customer service representatives, telemarketers, and stock and bond traders (Lee, 2017). Cenovus Energy Executive Vice President Kiron McFadyen, perhaps foreshadowing Suncor’s recent conversion of 400 heavy trucks in the Athabaska oil sands to driverless technology, used the terms “de-manning” and “zero-manning” the sector (Jaremko, 2017). AI is expected to outperform humans in many more vocations in the coming years. One study predicts that it will outperform translators by 2024, truck drivers by 2027, retail sales workers by 2031, fiction authors by 2049, and surgeons by 2053 (Grace, Salvatier, Dafoe, Zhang, and Evans (2018). Virtually all human jobs could be fully automated within a century, a phenomenon predicted by the economist John Maynard Keynes as permanent structural “technological unemployment” (Keynes, 1930).

While we have yet to see the full impacts of the current transformation, we can reasonably anticipate that these will be significant for society and the social sector. In these contemporary conversations about how we tackle the ‘social deficit’ and our seemingly intractable complex challenges and ‘wicked problems,’ it is unimaginable not to be thinking of the role of AI. We are starting to see the emergence of conversation starters on the topic like Ottawa-based Future of Good, think tanks like NESTA’s Centre for Collective Intelligence in the UK and the Center for Artificial Intelligence in Society at the University of Southern California. These relatively sparse discussions should be top of mind for social sector leaders, researchers and policy makers to generate a much deeper and far reaching dialogue that needs to occur.
With all these opportunities and risks at play, we need to imagine where society may be headed to consider how our sector can respond and be part of this AI-infused future. We have sketched four possible future scenarios, three of which take us in directions that are profoundly troubling for humanity. The fourth scenario, in contrast, is brimming with possibility.

One caution as well here: Futurecasting scenarios is a tricky business at the best of times. With a frontier issue such as AI, it is wildly moreso and prone to anthropomorphic bias, not to mention either naively wishful or dismally apocalyptic thinking (Yudkowsky, 2008). Most likely, AI will not play out in accordance with any of the following outlined scenarios. Rather, we will likely experience some kind of mélange of each scenario with a generous helping of the unforeseeable and unpredictable. Moreover, there is no single AI evolutionary path. AI is an incredibly diverse field and a host of innovations will emerge in different places and in different ways such that it will continue to resist sweeping generalizations (Ibid.).

The first possible future, though exceedingly bleak, is a future where civilization never comes close to realizing Artificial General Intelligence. There is ample evidence for the terra re-formatting collective destiny of our species, embodied in the geologic term the anthropocene, and documented in such recent publications as The Losing Earth (Rich, 2018), The Uninhabitable Earth (Wallace-Wells, 2019) and Bill McKibben’s (2019) Falter: Has the Human Game Played Itself Out? Climate change, soil depletion, loss of biodiversity and the continuing possibility nuclear conflagration may result in some form of civilizational collapse before we even have a chance to witness an AI ‘singularity.’ While plausible (Turner, 2014), this scenario does not take us any further in the AI and social good discussion, other than to say that early and mid-stage AI advancements may well prove to be indispensable tools in avoiding civilizational collapse.

Another possible (and still bleak) future is a civilizational threat directly resulting from AI. Technology entrepreneur and engineer Elon Musk warns that AI is the most important existential threat to humanity, referring to our pursuit of AI as "summoning the demon" (Musk, 2014). In this Terminator-esque scenario, Neuroscientist and philosopher Sam Harris maintains that AI will eventually be so exponentially superior to humans that we will be the equivalent of ants underfoot. He warns further that an AI arms race is almost inevitable (Harris, 2016). Nick Bostrum, head of Oxford University’s Future of Humanity Institute, authored a 2015 bestseller called Superintelligence: Paths, Dangers and Strategies. In it, he lays out a simple, but fraught dilemma: "As the fate of the gorillas now depends more on humans than on the species itself, so would the fate of humankind depend on the actions of the machine superintelligence." The book has been compared with Rachel Carson’s Silent Spring in its prescient, existential warning for the future of humanity. One online Amazon reviewer grimly referred to it as "an erudite book about how we all die." In a similarly foreboding article in The Atlantic entitled "How the Enlightenment Ends," former US Secretary of State Henry Kissinger warns that "philosophically, intellectually— in every way—human society is unprepared for the rise of artificial intelligence" (Kissinger, 2018). He notes AI’s difficulty in understanding context, in substituting wisdom for information, and in making rational inferences and pathways to solutions that may be difficult for humans to even comprehend. Never mistaken for a bleeding-heart peacenik, Kissinger goes even further in a new campaign to "Stop Killer Robots," pleading for a global ban on the use of autonomous, self-learning mechanized weapons systems (Knight, 2019).

The essence of yet another scenario, only modestly less bleak, was captured by the humanoid robot Android Dick in an interview for the PBS program Nova: "So don't worry," the robot reassured, "even if evolve into Terminator, I'll still be nice to you. I'll keep you warm and safe in my people zoo." In this "human zoo" scenario, the AI concludes that humans are either instrumentally useful (as in the film The Matrix), a nostalgic curiosity, or harmful such that our power must be thoroughly circumscribed. Think of HAL 9000 in the film 2001: A Space Odyssey, advising the human crew members that “this mission is too important for me to allow you to jeopardize it.” AI in this scenario will identify the flaws,
biases and crueler tendencies in humanity, concluding that our civilizational overreach is harmful to ecosystems and to ourselves as a species. As such, it will take steps to contain us, but – as its creators – it may also keep us ‘happy,’ superficially at least, in an innocuous setting where we cannot bring undo harm to others or to the machines. The algorithms may well ensure absolute equality (all humans are nourished and have equal opportunity for leisure, creativity and fun), but humans will lack agency, political power or any meaningful form of control. Some contend this is the best scenario we can realistically hope for.

However, a fourth future scenario, the one worth working on, accepts the inevitability of AI, but has us working vigorously to keep its development thoroughly in line with the very best of human values and aspirations. This is a human-machine co-created future. Some have even offered up a term for this: “Co-bots” (Daugherty and Wilson, 2018). As one futurist frames it, “Some humans will struggle against the AI. Others will ignore it. Both these approaches will prove disastrous, since when the AI will become more capable than human beings, both the strugglers and the ignorant will remain behind. Others will realize that the only way to success lies in collaboration with the computers. They will help computers learn and will direct their growth and learning“ (Roy, 2017). Such a future may not be utopia, but effective human-machine co-creation will surely help us eliminate homelessness, wrestle climate change to the ground, find cures and life extending treatments for countless diseases and put an end to war and violence as legitimate means of solving disputes. As Eliezer Yudkowsky, who popularized the term "Friendly AI" warns, "Artificial Intelligence could be the powerful solution to other existential risks, and by mistake we [may] ignore our best hope of survival” (Yudkowsky, 2008). Tech entrepreneur and futurist Jerry Kaplan argues that the darker AI scenarios — the idea of a runaway AGI confining or undermining humans — is far-fetched: He maintains a rosy optimism that we only ever use AI to solve human problems, presumably setting aside citizens as targets of marketing and surveillance (CBC Radio, 2019).

Another reason for optimism, for those in the caring and sharing professions, is that the most AI-proof jobs are those that require high levels of compassion, high levels of creativity, and especially both competencies in tandem: Caregivers, counsellors, teachers, artists, dramatists, etc. But in this scenario, we need to see these more as AI-ready, not merely AI-proof. Such helping professions and humanist vocations are uniquely positioned to step to the forefront of the debates on the future of humankind and can bring critical perspectives to the global transformation currently afoot; To help build what Lee (2018) calls a "blueprint for co-existence."

As Elizabeth Good Christopherson (2018), the President and CEO of the Rita Allen Foundation, observes, "Used poorly, there is no doubt that artificial intelligence can serve to automate bias and disconnection, rather than supporting community resiliency. For the social sector, a values-driven, human-centered, inclusive process of development can help to mitigate the ethical risks of developing artificial intelligence."

We tend to agree, perhaps overly optimistically, that a co-created future is possible. Plus, given the other possible future scenarios, what choice do we really have? The next question, then, is the role of the social sector and how do we prepare ourselves to be part of the creation of this future? What are the high leverage points and ecosystem conditions that we can influence in this discourse? And finally, how do we skill up for the challenges ahead?

*Developed by Hanson Robotics, it was never revealed whether the robot self-generated this reply or whether it was pre-programmed with this spine-tingling bon mot.*
To conceptualize how our sector can be part of a human-AI co-created future, we have to take a closer look at how AI is currently impacting key domains we are invested and involved in. Over the last two decades, most of the research and implementation of AI and Machine Learning has occurred in the domain of science more than any other. Beyond this, industries including transportation, marketing, and finance are particularly attuned to AI. It is only recently that these issues have drawn the attention of researchers, academia, organizations and governments interested in the application of these new technologies and advancements in the social arena. Arguably, humans know more about the social aspects of AI and Machine Learning through personal experience than through research.

Given the breadth of the social sector, it is helpful to cluster the impacts of AI further around the domains of health, the environment, arts and creativity, social good (the latter encompassing community development, social services, justice services, and human rights), and democracy.

**AI AND HEALTH**

Compared to other domains, we see some of the most significant advancements in the application of AI to healthcare. In the 1980s, academics were already foreseeing the impact that computers would have in medicine: As medical AI experts Dr. Glenn Rennels and Dr. Edward Shortliffe noted in 1987, "In time, computers may be as basic to medicine as the stethoscope. Medical systems can store data and retrieve it selectively; soon they will be ‘smart’ enough to advise in diagnosis and treatment" (Rennels & Shortliffe, 1987). AI innovator Neils Jacobstein has predicted 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" (Jacobstein, 2019).

The notion of failure to employ AI as a form of malpractice is a big leap from where AI’s relationship with healthcare started: There was considerable hype when IBM’s supercomputer Watson, which proved unbeatable as a Jeopardy! contestant, was repurposed to help in the fight against cancer. Watson has not proven to be as adept at oncology, with recommended interventions that have been more miss than hit. However, Alphabet's DeepMind, which has employed its neural networks technology in the service of medical diagnostics, has now matched the accuracy of medical specialists in analyzing 3D retinal scans and correctly diagnosing over 50 eye disorders. The technology is now in use at the Aravind Eye Hospital in Madurai, India, detecting diabetic retinopathy and diabetic macular edema, two major causes of blindness (Kelly, 2019). More recently, advances in the accurate diagnosis of pediatric diseases through AI, based on a sampling of over 1.3 million patient visits, showed comparable accuracy to experienced pediatricians (Liang, 2019). AI-enabled skin cancer detection is also outperforming dermatologists (Mar and Soyer, 2018).

‘Weak AI’ impacts on health care have long been transformative in fields such as anesthesiology, cardiovascular management, and procedure simulations (Rennels & Miller, 1988). Similarly, more recent machine learning systems have enabled systems that use enormous amounts of data gathered from patients in relation to chronic illness symptoms, which in turn enable diagnoses and treatment suggestions (Xu, et al., 2019; Gill-Cox, 2018; Zapusek, 2017; and Cannataro, Weber dos Santos, Sundnes, & Veltri, 2012). Fields such as radiology, pathology, ophthalmology, and cardiology have benefited from deep learning algorithms that have been able to diagnose diseases with a 96% accuracy rate, which has bested that of humans (Hsieh, 2017).

One study using machine learning, for example, showed tremendous promise in detecting signatures in the blood that could indicate the presence of an Alzheimer’s Disease marker otherwise only detectable in cerebrospinal fluid, at great cost and invasiveness (Goudey, Fung and Schreiber (2019). This could enable early treatment and dramatically improve the prognosis of those facing a future with Alzheimer’s. The algorithms employed in this study to identify proteins in the blood could be repurposed for
detecting certain other diseases in the same way.

A Stanford-based research project called Autism Glass detects emotions through facial expressions and translates these to cues the wearer can more easily interpret. In a randomized control trial of children with autism spectrum disorder, children who wore the device, based on Google Glass technology, showed a significant improvement in socialization over children receiving standard behavioural therapy (Voss, et al., 2019).

AI application in medicine has also reduced the administrative burden that medical professionals face in a daily basis: New tools for gathering information from patients, processing and analyzing results, diagnosing and treatment matching, monitoring and aftercare are facilitated increasingly by machine systems. Online resources, apps, bots and specialized software have improved the medical field and have resulted in a better patient experience (Mendeley, 2018). This can ultimately support vulnerable populations to have access to better medical help and treatment in more accessible and affordable ways.

Of course, we cannot overlook the impact of AI-assisted gene editing in the medical field and beyond. In November, 2018, He Jiankui shocked the world with his announcement of having created the world's first gene-edited babies (twins) using CRISPR technology – a powerful tool that enables scientists to alter DNA (Vidyasagar, 2018). This accomplishment was the result of years of research and development of the genomics field with the implementation of AI and machine learning in the process, allowing researchers to sequence and analyze DNA in more efficient and accurate ways (Marr, 2018).

With the advancement and funding expected for the genomics field in the near future, scientists will be able to better understand our genetics and make predictions about the likelihood of developing diseases, and as such develop personalized medicine for the specific needs of the population (Marr, 2018). In fact, the DNA of the twins born late in 2018, was edited to reduce their risk of contracting HIV. Such advancements in gene editing through AI and machine learning have already been applied in the agricultural and food industry to improve crops and engineer probiotic cultures (Marr, 2018). This crucial scientific development opens up the possibility not only to ‘optimize’ health but to ‘optimize’ the human race, along with the moral dilemmas that such power entails (Glasure, 2018).
AI AND ENVIRONMENTAL PROTECTION

The Fourth Industrial Revolution has been accompanied by redoubled and intensified efforts to embed sustainability in production and consumption, perhaps best exemplified in the United Nations’ Sustainable Development Goals (UN SDGs). The UN SDGs aim to challenge us globally to combat climate change, use resources wisely, develop sustainable cities and provide clean affordable energy (World Economic Forum, 2018). There has even been a Global Summit on the role of AI in advancing the UN Goals (Mead, 2018), and there are many people across the planet focused on how machine learning can lead to better outcomes for soil, air, water and climate, for ecosystems, and for endangered and threatened species. A McKinsey Global Institute discussion paper, entitled Notes from the AI Frontier: Applying artificial intelligence for social good catalogued 160 deep learning applied innovations across all 17 UN Sustainable Development Goals (Chui, et al., 2018).

To this end, Artificial Intelligence and machine learning can become essential tools in achieving these goals. AI can help optimize energy system forecasting, energy-efficient building management systems, and hyper-local weather forecasting for crop management. The McKinsey report noted AI-enabled applications in detecting illegal logging and poaching, and in detecting pipeline spills and tailing pond breaks (Chui, et al., 2018). Many of these applications are paired with rapidly advancing drone technology (Bondi, 2018). The adoption of electric cars and optimization of industrial machinery and manufacturing can help reduce carbon footprints while AI monitoring of ecosystems can reduce environmental degradation and minimize human impact on natural habitats. AI is also being used to develop clean-tech innovations, including solar panels, batteries, and materials and chemical compounds that can absorb pollution or conduct photo-synthesis (Knight, 2018).

Food security is one of the greatest challenges for humanity, with 821 million people facing chronic food deprivation in 2017 (FAO, IFAD, UNICEF, WFP, & WHO, 2018), while the world’s population is expected to increase by 2 billion by 2050 (Intel, n.d.). AI has become key in addressing such issues. Algorithms are currently being used in the agriculture industry for livestock, water, soil and crop management, yield prediction, disease detection, weed detection, crop quality, etc. (Liakos, Busato, Moshou, & Pearson, 2018). The agri-tech industry has developed advanced algorithms able to improve every aspect of the industry, reducing potential risks associated with the production of foods. For instance, we now have alerts sent to smartphones to warn farmers about wind changes, or with satellite images to inform farmers about possible pests landing on the crop (Intel, n.d.). Similarly, all the data collected through sensors is enabling a better understanding of the environment and its interactions with crop, soil, water, etc., allowing more accurate and realtime decision making (Liakos, Busato, Moshou, & Pearson, 2018).

When one considers an AGI that can digest more physics papers in a second than a human could in a thousand lifetimes, the prospects for new sources of energy, for example, are tantalizing to consider (Hodson, 2019). In the nearer term, narrower forms of AI have already made inroads into smart energy grids (helping enable a transition toward decentralized grids), as well as in urban transportation and smart agriculture (WEF, 2018). Optimizing and fine-tuning can result in significant efficiencies, and more precise climate modeling will help with climate adaptation and natural disaster response. DeepMind claims to have reduced sister company Google’s energy costs by 40% using algorithms that calculate the most efficient means to cool energy-intensive data centres.

There is an important question to consider in the mechanistic approach to the protection of nature: How does co-creation of AI comport with Indigenous and other holistic views of the natural world? Indigenous peoples have mobilized globally to support enhanced environmental protections in their homelands and traditional territories, while raising awareness of the need to respect our role and relationship with the land. The rights enshrined in one of the 21st Century’s most vital accords – the UN Declaration on the Rights of Indigenous Peoples – must inform and be embedded in AI design with respect to land, resource and other relevant decision-making. It is worth considering how multi-generational accountability and ecological carrying capacity limits can be embedded in AI, and how AI

---

Footnote: 5For a more detailed description of the current environmental challenges and the way AI can improve our relationship with the planet, see: World Economic Forum, 2018.
could assist in the delivery of ecosystem services, water keeping and land guardianship. As such, Indigenous knowledge holders must be involved in AI co-creation efforts, and at every stage.

One of the more ambitious and exciting uses of AI is the Earth Bank of Codes, an initiative from The Amazon Third Way, which aims to create an open, global platform that registers all forms of biological life and natural assets, including using blockchain to maintain spatial and temporal provenance, and codifying the rights and obligations outlined in the Nagoya Protocol of the Convention on Biological Diversity (WEF, 2018). A “biological search engine” will allow users to understand more fully the planet’s web of life, including the conservation status, ecological, biological, bio-chemical, biomimetic and traditional-knowledge attributes of every species, as well as the materials and chemical compounds that accompany each species. As an ecological research tool, this would prove invaluable.

**AI, ARTS & CREATIVITY**

When the first painting ever generated using AI emerged, a portrait entitled Edmond de Belamy, it created a stir, selling for US $432,500 at auction (Schneider & Rea, 2018). Beyond painting, AI is already being used in virtually every realm of the arts, from writing to music composition. Computational creativity, “the study of building software that exhibits behavior that would be deemed creative in humans,” has brought AI one step closer to fully emulating humanity. AI has been used to write poems (Robitzski, 2018) and scripts (Newitz, 2016), compose music (Metz, 2017), play poker (Riley, 2017), chess, checkers and Go (Metz, 2016), and generate images (Ha, Jongejan, & Johnson, 2017). One creative technologist has developed a machine learning homage to Jack Kerouac, in the form of an AI-generated novel called 1 the Road (CBC Radio, 2019).

The Expressive Machinery Lab at Georgia Tech has developed a technique called “collaborative movement improvisation,” where dance moves become data, taking cues from its human dance partners (Geggel, 2016). A machine learning technique called “generative adversarial networks,” which is the technology responsible for generating “deep fakes” (simulated photographs and video), also has incredible creative potential to aid in democratized, decentralized and dematerialized motion picture production. AI is predicted to write essays, mimic musicians, generate pop songs, produce creative videos, and even write a New York Times best seller by 2049 (Hall, 2018). A group of Stanford University researchers have created an AI application called Augur, that learns from a large corpus of amateur fiction to understand everyday human behaviour (sleeping, waking, eating, etc.) (Fast, et al., 2016).

We might be at the precipice of a cultural revolution triggered by the development of artificial intelligence. As Taichi Fukuyama, the creator of Amadeus Code notes, “History teaches us that emerging technology in music leads to an explosion of art (Hu, 2018). Common ways of conceiving art are being augmented and/or reinvented through algorithms challenging us to consider what it is exactly about us that makes us human and distinct from our machine creations. When almost every single aspect of our lives is being artificially enhanced and simulated by AI, it begs the question of what makes humans human, if not our intelligence, imagination and creativity. Visual arts (Rea, 2019), music composition (López de Mántaras, 2016), and even arts-based strategic foresight work (Plummer, 2017) are increasingly permeated by AI.

Perhaps one of the most astonishing advancements of artificial intelligence has to do with the development and sophistication of technologies such as Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR). These technologies allow us to combine our real world with a computer-generated world, giving us completely new experiences, insights and creative possibilities. AR layers digital information to the real environment using a camera or smartphone, as opposed to VR that generates a fully artificial environment (New Gen Apps, 2017). MR, in turn, combines both virtual and real elements of the environment creating scenarios where digital and real objects interact in real time (RealityTechnologies.com, n.d.).

Currently, AR is being used for 3D viewing in retail, e-commerce, and architecture, but also photography, cinematography, real estate, driving, camping, gardening, etc. Argon4 and ARLab display the contextual information of an image captured by a phone, or through games such as Pokémon-Go, Parallel Kingdom, Zombie Go, etc. (New Gen Apps, 2017). Similarly, VR is currently being used in numerous ways helping different industries to improve performance (Limbic Life Project VITALICS), improve marketing (Boursin Sensorium, TopShop Catwalk VR Experience, Toms Virtual Giving Trip and Adidas Delicatessen), offer training, (Lowe’s Holoroom) and raise awareness (Defy Ventures and Within: Step To The Line) (Becker, 2018). Microsoft’s HoloLens is probably the most notable example of MR development aiming to optimize businesses through Dynamics 365 applications (Microsoft, n.d.). In all, these ways of seeing the world through AI-lenses will undoubtedly impact our imaginations, worldviews and, ultimately, how our brains develop.
AI AND EDUCATION

Education at all levels will be disrupted by AI, but the benefits that AI brings are likely to far outweigh the upset in the system. A McKinsey Global Institute discussion paper notes, in particular, how AI is already enabling adaptive, student-centered learning applications (Chui, 2018). In particular, new technological devices and apps have allowed school teachers to develop strategies for personalized learning that was difficult to achieve due to large classroom sizes and short time to give to every student. AI and machine learning enable teachers to design/use apps and software to prepare materials according to every student’s need, in an understanding that the most effective model of instruction follows the model one student to one teacher. Moreover, through adaptive software, teachers can “assess, deliver content, and modify a student’s path through curriculum according to a set of programmed instructional heuristics” (Scharton, 2018).

Artificial intelligence offers an array of pedagogical tools to support K-12 students’ learning and teachers personalized learning attempts. Currently, a combination of augmented reality "smart glasses" for teachers to assess student progress and compare performance, cognitive tutors (a combination of computer science, cognitive science and big data to customize instruction), and mixed reality (platforms that combine the physical and the virtual world aiming to improve learning skills) is being used and tested in some schools (COSN, 2018). Similarly, AI’s advancement has enabled teachers to develop strategies to monitor students’ internet searches and even identify students at risk through machine learning algorithms. GoGuardian internet content monitoring is a software that prevents students from accessing damaging content and generates alerts to staff members about searches that can potentially lead to catastrophic events. Using this technology, members from a school in Florida were able to identify and provide appropriate intervention services to a child that had planned to commit suicide (Peterson, 2018).

Educational applications of AI are not confined to primary and secondary school. There have been many past dire warnings about the resilience of post-secondary education in the face of rapid technological change. Such warnings have also accompanied the rise of AI, but, this time, universities may not escape unscathed. Consider, for example, that AI can now generate academic papers. The Applied Computational Linguistics (ACoLi) lab at Goethe University in Frankfurt published an AI-authored textbook on the subject of lithium ion batteries, distilling insights from over 53,000 papers on the subject published in just the last three years (Beta Writer, 2019). The algorithm sorts and summarizes the peer-reviewed publications into coherent chapters and headings, turning hundreds of thousands of pages into a 180-page text. Published in 2019 by the science journal Springer Nature, this is the first machine-generated textbook in the world. The author credit is simply "Beta Writer."

Obvious though this may seem to someone outside the education system, the future of education will not be about filling up students’ minds with facts and definitions, as Northeastern University president Joseph Aoun (2018) lays out in his book Robot-Proof: Higher Education in the Age of Artificial Intelligence. Instead, he argues that the role of education must be to understand how we can use better data and machine learning, but also fill needs in society that even the most sophisticated AI agent cannot. 17-year-old 1st year university student Nhi Doan (2019), writing in a World Bank development blog, observes that “change is the only constant,” where students experience a world “so inundated with super intelligent machines, algorithms that can read our moods, and the constant remolding of jobs, [that] college education will have to be...reimagined.” 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.

\*Reconciliation efforts in places like South Africa, Rwanda, New Zealand and Canada include calls to reconcile our very different approaches to viewing, understanding and exploiting the land. Bolivia and Ecuador recognize constitutional rights to nature and establish punishment to those whose intentions are to destroy it. As certain South American indigenous groups contend the Sumak Kawsay or ‘wellbeing of all’ will never be achieved if this massive destruction of nature continues.
AI AND OTHER SOCIAL GOOD APPLICATIONS

While we can conceive of the impacts of AI on social issues, compared to other fields in which it is being applied, the use of new technologies to address social issues remains relatively underdeveloped. We are seeing very early experimentation and application of machine learning among non-profit organizations in the US, where analysis of open source data has, for example, enabled the reporting and rating of racially-motivated harassment from police officers (Suarez, 2017). Deep learning has also been used to identify "high-risk" texters, dramatically shortening the response time for crisis counselling (Suarez, 2017).

Given the impacts of AI and automation on employment across income levels (Su, 2018) it is ironic that we might use it to also mitigate the replacement of human labour with machines. For example, machine learning can be used to identify the best ways to reskill those left unemployed in the changing economy. It could help decision-makers analyze the job market and predict employability (Mewburn, Grant, Suominen, & Kizimchuk, 2018) (Hugo, 2018), identify future job losses and market gaps to anticipate labour force development needs. In turn, governments can incentivize particular careers and industries.

To address social challenges like poverty, satellite imagery and algorithms are being used to identify wealthy and poor regions in Africa (Big Cloud, n.d.). With this knowledge, policies and interventions can be targeted to the areas in most need. AI can be used to focus educational training or food system (Big Cloud, n.d.), and make predictions to help prevent future social issues using Big Data. Such predictor models have been deployed in the US to identify homeless persons likely to become high-cost users of public services (Toros & Flaming, 2019), or families that are at highest risk of homelessness (Turner, 2019). Other apps using AI and machine learning are being used to provide help to people experiencing homelessness or at high level of vulnerability (ibid.).

Another area that machine learning may prove immensely useful is via predictive algorithms connected to violent or suicidal behaviour. This would have been unthinkable a decade ago, and any attempts to do so would have been rightly dismissed as techno-naivety in the extreme. One experiment that looks at predictive factors for riots, lynchings and other mob violence in Liberia is producing rich and often counterintuitive insights, using the machine learning techniques of lasso, random forests, and neural networks (Blair, Blattman and Hartman, 2017). The non-profit Crisis Text Line, which analyzes millions of texts to predict suicidal behaviour, revealed that the word “ibuprofen” is “16 times more likely to predict the need for emergency aid than the word “suicide” (Christopherson, 2018). In turn, this insight has enabled a reshuffling of the queue and lives have been saved. Further analysis with the same technology showed that crisis workers were more successful when they employed improvised, adaptable responses instead of scripted therapeutic interventions. Ironically, the AI is telling crisis line workers that they need to be more human and less robotic. Other studies are testing complicated algorithms based on decades of accumulated insight: A meta-analysis (Franklin, et al., 2017) of a half century of research on risk factors for suicide suggests that incredibly complex algorithms incorporating hundreds of variables would be required to have any predictive value. Despite this complexity, there are teams of researchers at the Royal Mental Health Centre in Ottawa and at Florida State University using machine learning to analyze social media activity (in the former example) and anonymized patient records (in the latter). Beta versions of this technology show promise: In the latter study, 80% accuracy at predicting suicide attempts within 2 years, rising to 92% accuracy within one week (Anderssen, 2018).

Of course, the use of AI can come with significant challenges. In the 1990s, the US justice system introduced algorithms to help effective sentencing and determine whether a person awaiting trial should be sent to jail, released, granted parole or receive help in order to get their life back on track (Mbadiwe, 2018). Recent research on machine learning and judicial decision making has surfaced algorithm bias and profiling, leading to low rates of accuracy (Angwin, Larson, Mattu, & Kirchner, 2016; Danielle, Guo, & Kessler, 2017; Mbadiwe, 2018; Berk & Hyatt, 2015). In this sense, AI confirms social norms and biases rather than challenging them through evidence, because said evidence (data used to feed the AI) is biased itself. These biases and ethical concerns are at the heart of many AI and machine learning discussions.
Despite the proposed objectivity of machine learning analysis, research has proven biases exist (Mbadiwe, 2018). Much of the issues with such analyses come from the use of insufficient data or biased data being used to run predictions that are in turn used to guide policy and legislation. The impact of deliberate information selection that benefits particular groups has proven to be highly problematic in the case of political elections. Similarly, discrimination, racism and profiling often appear as unattended issues related to the use of machine learning in government⁸, justice, and employment. According to Millar et al., “The AI that is trained on biased data sets can entrench and proliferate those biases in its outputs, leading to discriminatory applications” (Millar, et al., 2018). Immigration advocates in Canada have raised concerns regarding possible human rights violations in the use of algorithms in the immigration and refugee application processing (Molnar, 2018).

As noted by Stark (2018), the development of machine learning systems and artificial intelligence should include not only computer scientists, but more importantly sociologists, anthropologists, lawyers, economists and historians in an attempt to better understand the effects and potential of AI on Canadian social good. The integration of “high tech” and social innovation efforts will be essential moving forward, given the interdependence of our species on a careful yet agile approach. It is interesting to consider the implications of certain universal rights informing the design of machine-assisted social programs, or the public policy optimization informed by deep learning of literature on promising interventions and/or welfare state supports. 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. AI in this light may well prove to be the worst nightmare of status quo politicians (or of status quo nonprofits, for that matter).

Indeed, capitalism itself will need to be severely augmented in order for social good outcomes to be better under an AI-dominated future, otherwise dominated by a handful of data oligopoly billionaires. As China-based legal scholar Feng Xiang (2019), puts it, “The more AI advances into a general-purpose technology that permeates every corner of life, the less sense it makes to allow it to remain in private hands that serve the interests of the few instead of the many. More than anything else, the inevitability of mass unemployment and the demand for universal welfare will drive the idea of socializing or nationalizing AI.” One hopes this logic doesn’t stop at China’s borders, but rather is fueled and shaped by open, democratic norms, which leads us to look at our final area of practical application for social good, AI and Democracy:

---

³Interesting results were found when used machine learning algorithms to the design of policies towards school enrollment and learning outcomes in India. See: Brockman, Fraker, McManus, & Buddy Shah (2019).
⁴Large amounts of data into Machine Learning systems have also been used to national security purposes and policing. However, the same discrimination and profiling issues have been evident in its implementation. For more on this, see: Stark (2018) and Leese (2014).
The convergence of politics, government and machine learning is one of the defining features of our times. Recently, the blending of politics and algorithms in incidents such as the Cambridge Analytica scandal have drawn the attention of the public due to evidence of campaigns using information about voters to influence election outcomes, in part through the spread of misleading facts. The precipitous drop in trust in nearly all institutions, tracked in Edelman’s Trust Barometer, was characterized in their 2018 edition as “The Battle for Truth” (Edelman, 2018). AI in the service of contaminating online discourse and amplifying vitriol is poison to the health of our democracies: As Filippo Menczer, Director of the Center for Complex Networks and Systems Research at Indiana University notes, “If people can be conned into jeopardizing our children’s lives, as they do when they opt out of immunizations, why not our democracy?” (Menczer, 2016). Though arguably the theme that is generating the most media attention, research on AI’s implication for democracy, is still in the early stages.

New technologies can paradoxically undermine and improve democracy (McGinnis, 2013). As one commentator puts it, “because algorithms (and yes, sometimes AI) are enmeshed in political decision-making, these technologies also offer a vision of ‘social good’ that can compete with liberal democratic commitments” (Shoker, 2019). Concepts such as “Digital Parties” (Chadwick & Stromer-Galley, 2016) or “Digital Democracy” (McGinnis, 2013) are being used to describe new modes of activism and political participation. The primary agora of the political debate is now social media. On the other hand, AI is being applied to emphasize the power of the state: the European Union (Fingas, 2018) and Canada (Wright, 2018) are probing the use of AI in screening of people at the port of entry to detect lies and possible national threats, sorting temporary visa applications (Wright, 2018), and granting refugee protection (Molnar, 2018).

It goes without saying that bias in the data can result in questionable decisions in the AI algorithm: AI can compound discriminatory sentencing, for instance, based on already skewed data. Significant concerns regarding AI and machine learning use to the benefit of elites have also emerged (Chadwick & Stromer-Galley, 2016) (Hughes, 2017). Yet, amid this depressing landscape, social good that can be realized, and might even flourish: As machine learning makes predictions based on large amounts of information, governments could greatly benefit from systems that allow the development of better policies based on predictions of positive outcomes. Governments could better foresee the impact that particular policies would have if implemented or run tests to obtain a description of the best policies to implement to solve specific social problems. If properly regulated by governments, AI and machine learning could greatly contribute to understand the impacts of global warming and pollution and help predict earthquakes or tsunamis. It could help governments to deploy better resources to deal with such situations in a more accurate and more affordable way.

It is also likely, as is happening already, that AI manifests itself simultaneously in very different ways in autocratic vs. democratic regimes, likely with dramatically different implications.

---

Some other examples of this could be found in: Bennett, Xiao, & Armstrong, 2004; Abernethy, Chojnacki, Farah, Schwartz, & Webb, 2018; Aras, 2008; and Iglesias, Mora, Martinez, & Fuertes, 2007.
With this technology revolution afoot, what role can we play and add value to as part of our contribution to positive social change within the Fourth Industrial Revolution? If we view the explosion of AI technology, both in ubiquity and sophistication, as just that, an "explosion," then we can extend the metaphor to see our role as helping carefully guide where the charges are laid. Nick Bostrom characterizes this as "the most careful controlled detonation ever conceived." How will the detonation be controlled such that our social foundations and ecosystems are maintained and optimized, while the biased, exploitive and inequitable implications of AI are avoided. As mathematician Cathy O’Neil points out, it’s not the algorithms themselves that are responsible for the socially destructive effects of machine learning – it’s the human bias already implanted in the coding of the algorithm; Algorithms are best thought of as "opinions embedded in math" (O’Neill, 2018). But which opinions have mattered so far? And from where do they originate? As we have already seen, the answer to this is currently far less pro-social than what we think should be the case.

While by no means exhaustive, we propose several competencies and conditions that those working in the social good sector need to get very serious about investing in over the coming years, within and outside civil society and the social good ecosystem. These competencies and conditions help us overcome fear of AI, while at the same time gaining agency and voice in how AI is developed. Each of the competencies and conditions extends beyond the individual to the role of civil society, the public sphere and even to the private sector, which is increasingly professing interest in social responsibility.

**1. An AI Commons**

Amy Webb (2019) notes that there are only nine companies, all based in either the US or China, that currently control the vast majority of AI development. These tech titans "are developing the frameworks, chipsets, and networks, funding the majority of research, earning the lion’s share of patents, and in the process mining our data in ways that aren't transparent or observable to us" (Webb, 2019). AI developers are in an out-and-out competition with each other, and the terrain of innovation is almost entirely on proprietary platforms. She argues further that this dynamic is the major factor influencing whether AI is socially beneficial or socially detrimental: "Safe, beneficial technology isn't the result of hope and happenstance. It is the product of courageous leadership and of dedicated, ongoing collaborations. But at the moment, the AI community is competitive and often working at cross-purposes" (Webb, 2019).

McKinsey Global’s analysis of barriers to AI for social good puts data accessibility at the top of their list (Chui, et al., 2018). In her 2019 Blueprint, Lucy Bernholz argues 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.” In a related blog post, she adds that policy innovators, software coders, and data collectors should assume that any AI “applied to an already unjust system will exacerbate the injustices, not magically overcome these systemic problems” (Bernholz, 2018). A data and technology commons would have many features, according to Bernholz, but would have to consider personal agency (including civil liberties and human rights), finding a more optimal balance between economic rights with public good imperatives (for example, around copyright law), privacy, structural issues like net neutrality, broadband access and ownership, how census and other public data is handled, and how state-market intersections are handled regarding surveillance, data storage and access.

Canada’s Digital Charter, announced in May 2019, is a positive step, committing to universal access: “All Canadians will have equal opportunity to participate in the digital world and the necessary tools to do so, including access, connectivity, literacy and skills” (GOC, 2019, May 21). The federal government has also mandated that the civil service make their source code open via the Open Resource Exchange (Shoker, 2019). Open and
collaborative data initiatives like Open North and Data for Good in Canada are at the vanguard of society’s impending battle for an AI commons. Funders take note: Watchdogs and data activists also need support. And we must protect, and honour, the data scientists and other employees who, for example, refuse to apply their skills toward the development of military applications (Shoker, 2019).

Although we have referred to a social sector digital divide, this is only partly true. While many local-based charities and other nonprofits fell behind the private sector in the tech race in the late 1990s and early 2000s, there was a parallel renaissance afoot in the form of open source coding (e.g. Linux), Peer2Peer file sharing, Creative Commons, and wikis, which are all concrete manifestations of a digital commons. Jimmy Wales’ commitment to maintaining Wikipedia as a nonprofit foundation is remarkable compared to Facebook, for example, which is not a digital commons. It is a proprietary space where your data is for sale, including to interests you may find repugnant.

**2. HYPER-CITIZENSHIP**

To be successful in this new era, both extreme enlightenment and hyper-citizenship skills are essential. Increasingly, the public is challenged to distinguish between the fake and the real: to know whether what is being seen and heard is authentic or manipulated. The line between objectivity and subjectivity is thinner than ever before, and as George Orwell put it in his book 1984, “The very concept of objective truth is fading out of the world. Lies will pass into history” (Orwell, 1949). By 2020, some analysts believe that “AI-driven creation of ‘counterfeit reality’, or fake content, will outpace AI’s ability to detect it” (Columbus, 2017). We can only expect to see new and more sophisticated forms of digital skullduggery, like the ‘deep fake’ or other strategies of manipulation of audio and video to create artificial speeches and embarrassing or incriminating scenarios that are false but assumed to be real, with catastrophic impacts (Adler, 2017).

In this light, a skeptical society and critical thinking skills are more necessary than ever. Knowing ‘real’ truth will be more challenging and time consuming than in the past given the rapidly growing sophistication of manipulation, but such vigilant sleuthing and skepticism can make us better citizens.

AI may also help us find a pathway to rational compassion rather than relying on our imperfect, hyper-biased sense of empathy. 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 (Bloom, 2016).

---

9Described and demonstrated by tech futurist Simon Adler in his interview with the podcast RadioLab: http://www.radiolab.org/story/breaking-news/
With jobs in manufacturing, sales, transportation, accounting, security, diagnosis and basic research and storytelling (including much of journalism) are likely to disappear in the not-so-distant future. However, careers involving caring, complexity and creativity will continue to exist and our imagination will be better nurtured and rewarded than previous industrial eras. We may even bear witness to a post-digital Renaissance. Canadian futurist Hamoon Ekhtiari (2018) argues that to approach AI productively, we need not just an ethical framework and social purpose orientation, in addition to being really, really smart and plugged-in. He argues that we desperately need to foster a collective imagination in order for AI to truly serve humanity. Marshall McLuhan said that art is a "distant early warning system." As such, artists need to be at the centre of AI development. New Yorker writer Tad Friend further predicts that we will need "experts in unexpected disciplines such as human conversation, dialogue, humor, poetry, and empathy." As such, the humanities will take on a renewed relevance and import, in the service of fostering creative mindsets, systems thinking and mental elasticity.

Philosophers, for the first time in history, might find themselves having an in-demand, marketable skill set. The ethical frameworks for AI might benefit, from a social justice standpoint, from embedding such notions as Peter Singer's "effective altruism" (ensuring global fairness in the relief of poverty), to John Rawl's "veil of ignorance" (ensuring genuine equality of opportunity and eliminating barriers to social mobility)\textsuperscript{11}, to Susan Moller Okin's feminist modifications to liberal theories of justice, as just a few potential examples.

The social sector may withstand, at least in theory, the impacts of AI and automation because of its focus on caring at the frontline levels. However, we cannot assume the way we work now will remain remotely the same. Already, we see the introduction of low-cost counselling services using online platforms, the disruption of Open Data for nonprofit/charity/government funding models, and the democratization of information using technology to challenge traditional service access pathways. This is not to lament these changes; In fact, we can make the case from the user perspective that technology has empowered consumers with real time information about benefits and services they can access. Similarly, the donor and taxpayer can now understand the financial flows into the social safety net in a much more transparent fashion.

These are opportunities for creativity and rethinking of complexity, leveraging AI to support social impact in ways we could have never previously imagined. AI’s potential to amplify our talent, creativity, openness, and capacity for inclusion should make all people who share and care for a living incalculably more effective and valued.

\textsuperscript{11}See, for example, Eriksson, K. (2018). Realizing Rawls in an Automated Future (Graduate Thesis). Lund University, Sweden.
Technological and data literacy will be essential for the social sector to both support better machine learning analysis and to develop better policies and societal outcomes. But, as Alix Dunn, founder of The Engine Room, which connects data and tech to social impact work, reassures us, "Very few of us feel we have the knowledge and confidence necessary to direct and shape it towards specific ends. Fear not! The truth is we don’t all need to learn to code to navigate and influence our digital world" (SSIR, 2019). That said, it is also true that we have a lack of talent in the pipeline specific to socially-oriented AI development (Chui, et al., 2018). And competition for talent for the foreseeable future will be fierce. The nonprofit sector is historically on the losing end of such battles, although it will help that we see more and more computer science, engineering and business students embrace social impact work.

Karen Hao, MIT Technological Review’s AI reporter, contends as well that "we need to stop perpetuating the false dichotomy between technology and the humanities" (Hao, 2019). She argues that, in order to build more ethical products and platforms, software engineers and programmers need better grounding in the liberal arts. Conversely, policymakers, social change-makers and civic leaders need better technology literacy. AI bootcamps and short intensives for social sector managers, designers and evaluators could prove useful. Universities should consider social impact work-integrated residencies for AI specialists finishing their PhD or other advanced credentials (Chui, et al., 2018). Without integration of these skills, we risk becoming marginalized from both debates and practical applications of these new technologies. As Kissinger points out, "Philosophers and others in the field of the humanities who helped shape previous concepts of world order tend to be disadvantaged, lacking knowledge of AI’s mechanisms or being overawed by its capacities" (Kissinger, 2018).

There is extreme risk in our attempts at social innovation being merely social engineering, with unforeseen detritus littering the wake of good intentions. We have seen what happens when we leave affordable housing solutions to architects, or city-building to transportation engineers. Similarly, leaving social application development to computer scientists and data specialists is a recipe for guaranteed unintended harm, no matter how well-meaning the aim. Enhanced efforts to integrate social innovation and high-tech innovation will be essential moving forward, including working as diverse, interdisciplinary, cross-functional teams who can help anticipate and mitigate bias, assumptions and unintended consequences. As such, people who care about and know a lot about (including having lived experience with) a given social issue must be deeply embedded at every stage of machine-enabled deep learning regarding that issue. We also need “data translators” in the sector – NGOs employing data scientists who can interpret and stress-test an algorithm’s “brittleness” and bias vulnerability (Chui, et al., 2018). Nick Bostrum (2015) adds that as we create a really powerful optimization processes to maximize for a given social objective — call it Objective X — we had better sure that our definition of ‘X’ incorporates everything we care about. It should also incorporate the perspectives of citizens/clients/users/patients — whomever might be at the ‘receiving’ end of an AI application intended to improve a human or social service.

While the private sector, and, to a lesser extent, the government have enjoyed access to technology software and hardware, the nonprofit sector has always faced a significant challenge accessing basic technology infrastructure. It is not uncommon that community organizations run on donated or low-cost machines, with obsolete software and slow performance. Tight resources have made investing in new technologies, software, IT and professional development on AI difficult as well. This has hampered capacity development in the frontline and management layers of service providers’ learning about these technologies and how they could benefit the population they serve. Encouragingly, as tech gets easier to use and cheaper to buy, this divide is closing. Funders can also play an obvious role here to help close the gap. US-based initiatives such as the Partnership for AI and AI4ALL aim to make AI approachable to the general public, recognizing that AI can be an intimidating topic, and those who engage with the topic are disproportionately male, privileged and working in the commercial sector. Only 10% of Google’s employees working on "machine intelligence," for example, are female (Wykstra, 2019). Google has also invested in the STEM education organization Actua to develop an artificial intelligence curriculum for high school students across Canada (Shekar, 2019).
Participating in the co-creation of a future with AI will require our sector to embrace risk and have a stronger voice in public decision-making. We will need to adapt to and influence future technological advances while developing better services and contributing to social justice goals. We would be foolish to assume we are disruption-proof. Take as one small example the emergence of Benevity, which has disrupted workplace employee giving such that it is displacing the United Way campaign model in some Canadian cities.

We will need to develop a different lens for risk internally to fully take advantage of the opportunities ahead. This should be a wake-up call to nonprofit boards, many of whom are notoriously risk-averse. An important dimension of this is data sharing between organizations: “Open data” has ruptured barriers that we took for granted. There are many other areas in which we will need to adapt and pursue social impact very differently in such contexts at the frontline, policy and funding levels. The age of charity may be yielding to an age of shared, collaborative social infrastructure (Stauch and Johnson, 2018). Many organizations will not survive these transitions.

As Kai Fu Lee argues, “AI is serendipity... It is here to liberate us from routine jobs, and it is here to remind us what it is that makes us human” (Lee, 2018). With less people needed to perform routine tasks our sector will find itself caught up in the broader shifts in work, as well as in attitudes toward work and living. With more time available to dedicate to creative and recreational pursuits, and to connect with nature, we may expect the arts, sport, environmental protection and entrepreneurship to flourish. Citizens could become more responsible, critical and skeptical about the information they receive. Governments may also have a more diverse array of potential representatives. Our ability to innovate and adjust to this new reality can contribute to a constructive unfolding of this new industrial and technological revolution, but only if we are leaders and active participants in this change.

5. DISRUPTION TOLERANCE
The visibility of AI’s social implications and transformational possibilities must be greatly elevated in the public sphere. If the monopoly power of commercial tech giants and totalitarian regimes is not greatly circumscribed, AI will fail to serve the common good. At a global level, we are seeing the emergence of concords such as the Montreal Declaration for a responsible development of AI, “Born from an inclusive deliberation process that initiates a dialogue between citizens, experts, public officials, industry stakeholders, civil organizations and professional associations” (Montreal Declaration, 2018). Emphatically voluntary and aspirational, there is a need to build on such initiatives with anticipatory regulation, legislative commitments and multi-lateral teeth, as we see happening with the Ministerial Declaration on AI in the Nordic-Baltic Region. Facilitated through the Swedish-based Future of Life Institute, the policy objective is to develop and promote the use of AI “to serve humans better.” A timely initiative has been launched in Canada by the Canadian Institutes for Advanced Research and the Brookfield Institute for Innovation + Entrepreneurship, bringing policy innovators from the public, private, academic, and not-for-profit sectors in a series of conversations about the regulatory and other public policy implications of AI (Villaneuve, Boskovic and Barron, 2018).

In December 2018, Canada and France announced the creation of an International Panel on Artificial Intelligence (IPAI) which is to include representation from civil society and will align AI investment to the UN SDGs (GOC, 2019, May 16). Despite this commitment, in international arenas Canada has been weak in its criticism of autonomous weapons systems and other forms of socially malignant AI (Shoker, 2019).

Such global and regional efforts also need to penetrate the public consciousness within nations and cities. Existing or future networks and coalitions of nonprofits, foundations, and social innovation practitioners should similarly be keen to enhance awareness and visibility of AI, and be active agents in the promotion of responsible AI. This is one issue that the nonprofit sector should not be deferential to. Social purpose organizations cannot be at the AI “kids table” hoping to be asked to sit at the big table. There will be those who use techno-obfuscation to keep the social purpose voices on the margins, but there are likely many more allies and champions who would welcome a strong social sector voice in the development of accords, protocols and processes.

6. PUBLIC COMMITMENTS, DECLARATIONS AND PROTOCOLS
AN OPPORTUNITY FOR WORLD-LEADING INTEGRATION OF SOCIAL INNOVATION AND AI DEVELOPMENT

Canada has a world-leading record in AI science and tech innovation. We are also increasingly recognized globally as a powerhouse in social innovation. Our collective impact work, social innovation education and practical innovations in public policy are places where Canada is standing remarkably tall. But social innovation is not tech innovation, and we have so far failed to bring those worlds together effectively. People working on social innovation occasionally consider the role of tech, and people working in tech innovation occasionally are motivated by social purpose. But too often, these worlds are speaking completely different languages and talking past each other. As Bernholz (2018) 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 its past time to keep pretending they will.” One point of divergence between the tech and social innovation worlds, among many others: Are we promoting commercialization of intellectual property or are we working to strengthen the commons?

To help bridge this divide, we need mediated spaces – from formal think tanks and university-based centres, to book clubs and salons – where the relationships, trust and shared platforms for ideas and experiments can flourish, and where radical, even revolutionary, action can emerge. In addition to the new centres mentioned previously at UGuelph, ULaval and UToronto, we should take inspiration from leading global centres such as the Centre for the Study of Existential Risk at Cambridge, the Future of Humanity Institute at Oxford and the Swedish Future of Life Institute, all thinking deeply about the role of tech and the future of our species. We should further seek to create and support initiatives that zero-in specifically on AI and professional ethics, as the American Society for the Advancement of AI has done since its founding in 1979. Other institutions that focus on AI ethics include the AI Now Institute at New York University, the AI Ethics Lab, the Algorithmic Justice League, Safe AI, the Open Roboethics Institute and the Foundation for Responsible Robotics. Beyond ethics, we need concerted, sustained and well-resourced efforts to look at the broader social purpose potentials and implications of AI. Inspiration for this can be found at Cambridge University’s Leverhulme Centre for the Future of Intelligence, the Centre for Artificial Intelligence in Society at the University of Southern California, the UK-based Centre for Collective Intelligence Design (a Nesta initiative) and MIRI Berkeley, the Machine Intelligence Research Institute. Existing “social R&D” centres and platforms, though they are few and sparse on the Canadian landscape, could serve as loci for conversations, experimentation and knowledge-bridging.

Incentives and prizes to bring AI and social good together are also needed. Google.org provides a $25 million prize to initiatives using AI to help address social challenges. Their DeepMind subsidiary operates research programs in Ethics & Society and Health. The World Bank, AI for Good Foundation and the Rockefeller Foundation’s Datakind program also provide funding to bridge AI to social challenges. In Canada, we likewise need funders and investors, from venture philanthropists to government research programs, to promote AI and common good work. We also need more shared conversations, more scenario planning, and more strategic foresight. The Civil Society Futures project in the UK may offer some clues as to how to proceed. While the long-term impacts of AI and machine learning are yet to be fully understood, free online courses and machine learning software, as well as university programs focused on AI provide opportunities for social impact organizations in the private, NGO, and public sector to dive into learning about these new realities. There are many places one can begin their AI literacy journey.
As the co-founder of the Machine Intelligence Research Institute, Eliezer Yudkowsky, frames our current predicament, "Pragmatically speaking, our alternatives boil down to becoming smarter or becoming extinct... We must execute the creation of Artificial Intelligence as the exact application of an exact art" (Yudkowsky, 2008).

The careful combination of machine super-intelligence with human learning has significant potential to help us solve "wicked problems" (Hanna, et al., 2016). Machines will help us immensely with understanding complex social and ecological challenges (Mar, 2017). However, as Lucy Bernholz (2019) challenges us, this means that those who work on social issues must not simply engage with technology in a deeper way. It also means that we must re-imagine civil society itself:

"...how[might] a revisioning of civil society proceed when the starting assumption is that our individual and collective dependence on digital infrastructure will continue and get more complicated?"

Some of us were fortunate to benefit from a liberal arts education - the kind everyone (especially our parents) thought would never lead to a ‘real’ job. But it is this background that allows us bounce across disciplines that has served us well - we learned how to learn, which means we can learn anything. As one student puts it, "How else will we ‘run faster than the algorithms, faster than Amazon and the government, if not by living a life of perennial learning?’" (Doan, 2019).

We live in a time where the rebirth of the "Renaissance" man/woman/+ is afoot. Creativity is not limited to the arts; social purpose is not limited to charity; innovation is not limited to the tech sector; and profit is not limited to industry. Those who are able to take risks, engage curiously and critically with data and technology, and bend social purpose into something truly transformative will win the day. The question is, is civil society future-proof, and more importantly, what role will it play in the Fourth Industrial Revolution.

If this sounds like a lot of grueling work, that’s because it is. It is exponentially greater in scope, cost and risk than the Manhattan Project or the Marshall Plan. But thankfully there are nearly 8 billion of us who can share the burden. The stakes are stratospheric, but the rewards could be incalculable – the survival and thriving of the human family, within a restored and flourishing web of life. We must pour generous, never-yielding creativity, inclusion and rational compassion into this most important invention we will create as a human species. As Bostrum (2015) predicts, assuming we optimize for our long-term prospects by ensuring the right controls are in place for the first stages of super-intelligent AI, "Millennia from now, people will look back and note that the one thing we did that really mattered, we got right."

Former world champion chess player Gary Kasparov, defeated by an early AI in the form of Deep Blue, reflects that as computers focus on the dull, dangerous, dirty and dear, this should permit humans to elevate our cognition "toward creativity, curiosity, beauty, and joy." Kai Fu Lee (2018) offers yet another extraordinary coda to consider: AI may ironically remind us why we exist in the first place: It is not for work. It is for love.
James Stauch is the Director of the Institute for Community Prosperity at Mount Royal University where he has led the creation of non-credit, co-curricular and credit-based programs for students and practitioners in social innovation, community investment and the economics of social change. The Institute also co-developed the Trico Changemakers Studio, an on-campus community co-working and social R&D space. James previously served as a philanthropy executive and consultant, including as senior executive for the Walter & Duncan Gordon Foundation in Toronto. He has served as chair of the Arctic Funders Collaborative, International Funders for Indigenous Peoples, Circle on Philanthropy and Aboriginal Peoples in Canada, and the Canadian Environmental Grantmakers Network. James has published in American Behavioural Scientist, Urban Affairs, the Association for Nonprofit and Social Economy Research Journal, The Philanthropist, the Northern Review, and Voluntas. He is the lead author of an annual scan of trends and issues, produced in partnership with Calgary Foundation, and is a leadership faculty member for the Conference Board of Canada’s Corporate Responsibility and Sustainability Institute. James is also a regular contributor to the Future of Good and KCI Philanthropy Trends. His most recent publications include co-authoring a Students’ Guide to Mapping a System, published by the Skoll Centre for Social Entrepreneurship at Oxford University, The Loney Companion: 10 Steps to Starting a Social Enterprise in Canada, produced with Encompass Co-op, and a chapter on the role of business in the community, as part of a forthcoming textbook on the nonprofit sector in Canada, produced by Carleton University and the Muttart Foundation.

Dr. Alina Turner is recognized as a leading Canadian researcher and thinker on homelessness and her work on system planning is recognized as a leading practice and often called upon as a model internationally. She is a Fellow at The School of Public Policy, University of Calgary where she focuses academic research and publishing on systems integration to improve health and social outcomes. Alina leads Turner Strategies, a consulting firm that builds capacity in non-profits, government and private sector partners to accelerate social impact by leveraging research, community engagement, and creative technologies. She also co-founded HelpSeeker as a social enterprise dedicated to connecting people with the help they need, fast. The back-end of HelpSeeker provides service providers and decision-makers with real-time analytics to inform strategy and decision-making. HelpSeeker is fast becoming a system mapping tool across Canadian communities, essential to mapping and analyzing over 100,000 social and health resources nationally. Alina brings expertise in developing and implementing solutions to complex urban social issues like poverty, homelessness, mental health, addictions, and violence. She leverages research, technology, community engagement, and systems thinking to accelerate large-scale social change.

Camilo Camacho Escamilla is a political scientist with a Master's Degree in Public Policy from the University of Calgary. He has more than 5 years of experience in social research and policy analysis in diverse topics such as immigration, development, gender, integration, among others. He is currently a member of the Immigrant Advisory Table (IAT) of the City of Calgary and a member of the Board of Directors of the Association of Colombian – Canadian Professionals of Alberta (ACCPA). He is working with Dr. Alina Turner on the rollout of the HelpSeeker platform to support the transformation of Canada’s social safety net as Manager, Research & Policy.
The Institute for Community Prosperity, driven by MRU’s mandate to provide extraordinary opportunities for undergraduates, works to ensure that students and other citizens have access to high-impact, immersive, and uncompromisingly current learning to improve and transform communities; unlocking their potential, and helping them flourish as learners, changemakers, and human beings. The Institute mobilizes knowledge in the form of useable, insightful publications, and designs and delivers community-based and co-curricular learning.

MRU is proud to be an AshokaU designated Changemaker Campus.

ACKNOWLEDGEMENTS

This report was edited by Mason Benning, Journalism student at MRU, and designed by Elle Griffin, Information Design graduate at MRU.


social trust. From CIFAR. Retrieved from https://www.cifar.ca/docs/default-source/ai-society/g7-accountabilityinai.pdf?sfvrsn=66bdfddc2_6


