Economics
The Open-Access, Open-Assessment E-Journal

Building on the Hamburg Statement and the G20 Roadmap for Digitalization – Toward a G20 framework for artificial intelligence in the workplace

Paul Twomey

Abstract

Building on the 2017 Hamburg Statement and the G20 Roadmap for Digitalization, this paper recommends a G20 framework for artificial intelligence in the workplace. It proposes high level principles for such a framework for G-20 governments to enable the smoother, internationally broader and more socially acceptable introduction of big data and AI. The principles are dedicated to the work space. It summarises the main issues behind the framework principles. It also suggests two paths towards adoption of a G-20 framework for artificial intelligence in the workplace.

(Submitted as Global Solutions Paper)

Authors

Paul Twomey, Distinguished Fellow, Centre for International Governance Innovation, paul@stashdaddy.com

1   Introduction

In their declaration following meetings in Hamburg, Germany, in July 2017, leaders of the Group of Twenty (G20) countries formally recognized that "digital transformation is a driving force of global, innovative, inclusive and sustainable growth" and committed "to foster favourable conditions for the development of the digital economy and recognise the need to ensure effective competition to foster investment and innovation" (G20 2017c, 5).

With the theme of "shaping an interconnected world," the leaders also recognized that the swift adoption of information and communications technology is rapidly changing the workplace and placing stresses on citizens, societies and economies: "Well-functioning labour markets contribute to inclusive and cohesive societies and resilient economies. Digitalisation offers the opportunity for creating new and better jobs, while at the same time raising challenges regarding skills, social protection and job quality…Acknowledging the increasing diversity of employment, we will assess its impact on social protection and working conditions and continue to monitor global trends, including the impact of new technologies, demographic transition, globalisation and changing working relationships on labour markets. We will promote decent work opportunities during the transition of the labour market" (G20 2017c, 6).

Responding to the rise of big data — the explosion of data and datification — and artificial intelligence (AI)[1] is one of the most important ways that G20 leaders could address the goals in the Hamburg declaration at forthcoming meetings in Buenos Aires.

In Hamburg, leaders stated that "the G20 Roadmap for Digitalisation will help us guide our future work" (G20 2017c, 6). In that roadmap, a paper outlining policies and annexed to their declaration of April 2017, ministers responsible for the digital economy said that they would further discuss "frameworks as enablers for…workforce digitalisation" (G20 2017a, 7). They indicated what some of the aspects of such frameworks would be: "In order to better prepare our citizens for the opportunities and challenges of globalisation and the digital revolution we need to ensure that everyone can benefit and adapt to new occupations and skills needs…Trust and security are fundamental to the functioning of the digital economy; without them, uptake of digital technologies may be limited, undermining an important source of potential growth and social progress...Within the [2018] Argentinian Presidency of the G20 we will discuss international public policy issues related to privacy and security in the digital economy" (G20 2017b, 13).

These issues — trust, security, the need to adapt, privacy, skills — are central as workers and citizens react to the rapid introduction of big data collection and related AI. Confronted with

_____

[1] Nearly all software programs contain some form of algorithm and pose little disruption to the workplace. However, the complex algorithms that drive significant decision making in the workplace have drawn public attention. In this paper, AI refers to automated decision making informed by complex algorithms and machine learning capabilities.

forecasts that these technologies may affect nearly half of all jobs,[2] workers worry about their employment and what skills they will need. People seek assurance that AI and automation will be introduced in a manner ensuring respect for the human integrity of workers and under a framework of accountability, while still delivering the productivity, safety and innovation benefits promised.

This paper offers such a framework for G20 governments, to enable the smoother and more socially acceptable introduction of big data and AI. It explores the main issues involved and suggests principles for a framework. It also outlines two paths toward adoption of this proposed G20 framework for AI in the workplace.

2   The Issues

The use of automated decision making informed by algorithms is penetrating the modern workplace, and broader society, at a rapid rate. In ways not visible to, nor fully apprehended by, the vast majority of the population, algorithms are determining our present rights and future opportunities. To consider just one aspect of everyday life, automobile transportation, these algorithms help us drive our cars, determine whether we can get a loan to buy them, decide which roads should be repaired, identify if we have broken the rules of the road and even determine whether we should be imprisoned if we have (see Angwin et al. 2016).

2.1   Benefits

Big data and AI can provide many benefits. They can assemble and consider more data points than humans can incorporate and often provide less biased or clearer outcomes than humans making decisions. Examples include the prevention of medical errors to increasing productivity and reducing risks in the workplace. Even in the explicitly human function of the human resources department, machine learning can improve job descriptions and provide more "blind" recruitment processes, which can both increase the pool of qualified candidates and boost recruitment of non-conventional applicants.[3] Written well, algorithms can be more impartial and pick up patterns people may miss, in this and other applications.

Many commentators point to the productivity benefits of AI. For instance, analysis by Accenture of 12 developed economies indicates that AI could double annual economic growth rates in 2035: "The impact of AI technologies on business is projected to increase labor

––––––––––––––––––––––––––––

[2] For example, KPMG International (2016, 2) reports that "between now and 2025, up to two-thirds of the US$9 trillion knowledge worker marketplace may be affected. The Bank of England estimates that robotic automation will eliminate 15 million jobs from the United Kingdom economy in the next 20 years. Digital technologies will conceivably offset the jobs of 130 million knowledge workers — or 47 percent of total US employment — by 2025. Across the [Organisation for Economic Co-operation and Development], some 57 percent of jobs are threatened. In China, that number soars to 77 percent."

[3] See firms like Textio (https://www.textio.com/) and Pymetrics (https://www.pymetrics.com).

productivity by up to 40 percent and enable people to make more efficient use of their time" (Purdy and Daugherty 2016). The World Bank is exploring the benefits of AI for development and in uses from predicting migration patterns to reducing poverty.[4] Others identify farming, resource provision and health care as sectors in the developing economies that will benefit greatly from the application of AI (see Ovenden 2016).

2.2   Impact on Employment

Much has been made of the impact of AI and related robotics on jobs, especially since Carl Benedikt Frey and Michael A. Osborne's 2013 paper estimating that 47 percent of jobs in the United States were "at risk" of being automated in the next 20 years. Debate has ensued on the exact nature of this impact: the full or partial erosion of existing job tasks, the impacts across sectors and across developed, emerging and developing economies. Forecasting such effects is inherently difficult. But a recent summary from the McKinsey Global Institute reflects a midway analysis.

> Automation technologies including artificial intelligence and robotics will generate significant benefits for users, businesses, and economies, lifting productivity and economic growth. The extent to which these technologies displace workers will depend on the pace of their development and adoption, economic growth, and growth in demand for work. Even as it causes declines in some occupations, automation will change many more — 60 percent of occupations have at least 30 percent of constituent work activities that could be automated. It will also create new occupations that do not exist today, much as technologies of the past have done…

> Our scenarios across 46 countries suggest that between almost zero and one-third of work activities could be displaced by 2030, with a midpoint of 15 percent. The proportion varies widely across countries, with advanced economies more affected by automation than developing ones, reflecting higher wage rates and thus economic incentives to automate….

> Even if there is enough work to ensure full employment by 2030, major transitions lie ahead that could match or even exceed the scale of historical shifts out of agriculture and manufacturing. Our scenarios suggest that by 2030, 75 million to 375 million workers (3 to 14 percent of the global workforce) will need to switch occupational categories. Moreover, all workers will need to adapt, as their occupations evolve alongside increasingly capable machines. (Manyika et al. 2017, vi)

Whatever the specifics, the results are clearly going to be very significant for G20 economies and their citizens. And, if the rate of adoption continues to outpace previous major technological

_____

[4] See www.measuredev.org/.

adoptions,[5] the scale of social dislocation is likely to be greater — which provides even more reason for the G20 to work now on a framework for AI adoption.

2.3   Risk of Bias

Code is written by humans and its complexity can accentuate the flaws humans naturally bring to any task.

Bias in the writing of algorithms, as a product of human endeavour, is inevitable, and can have chilling effects on individual rights, choices and the application of worker and consumer protections. Algorithms incorporate built-in values and serve business models, which may lead to unintended biases, discrimination or economic harm.[6] Compounding this problem is the fact that algorithms are often written by relatively inexperienced programmers who may not have a correct picture of the entire application or a broad experience of a complex world. The dependency of the workplace on algorithms imparts tremendous power to those who write them. These programmers may not even be aware of this power or the potential harm that an incorrectly coded algorithm could do. Researchers have discovered bias in the algorithms for systems used for university admissions, human resources, credit ratings, banking, child support systems, social security systems and more. Because the complex market of interacting algorithms continues to evolve, it is also likely that existing algorithms that may have been innocuous yesterday will have significant impact tomorrow.

AI is subject to two significant types of bias:

- bias in its coding, or
- selection bias in or distortion/corruption of its data inputs.

Either type can result in significantly flawed results delivered under the patina of "independent" automated decision making.

2.4   The Criticality of Truly Applicable and Accurate Data Inputs

While much contemporary commentary has focused on the question of bias, the long experience of software development teaches that the proper scope, understanding and accuracy of data have dominant impacts on the efficacy of programming. In simple terms, "garbage in, garbage out." This relationship is particularly true with AI. AI is a process of machine learning — or, more accurately, machine teaching. The inaccuracies in data often come from reflections of human biases or human judgments about what data sets tell us. The establishment of training data and training features is at the heart of AI. As Rahul Barghava (2017) says, "In machine learning, the

_____

[5] See discussion in Lohr (2017).

[6] For instance, media reports (see, for example, Wexler 2017) have pointed out clear racial bias resulting from reliance on sentencing algorithms used by many US courts.

questions that matter are 'what is the textbook' and 'who is the teacher.'" The more scrutiny these can receive, the more likely that the data will be fit for purpose. To consider one example, some local governments in the United States have been making more use of algorithmic tools to guide responses to potential cases of children at risk. Some of the best implementations involve widespread academic and community scrutiny on their purpose, process and data. The evidence is that these systems can be more comprehensive and objective than the different biases people display when making high-stress screenings. But even then, the data accuracy problem emerges: "It is a conundrum. All of the data on which the algorithm is based is biased. Black children are, relatively speaking, over-surveilled in our systems, and white children are under-surveilled. Who we investigate is not a function of who abuses. It's a function of who gets reported."[7] Sometimes the data is just flawed. But the more scrutiny it receives, the better it is understood. In the workplace, workers often have the customer and workflow experience to help identify such data accuracy challenges.

Acceptance of data inputs to AI in the workplace is not just a question of ensuring accuracy and fit for purpose. It is also one of transparency and proportionality.
The recent crisis surrounding Facebook, over Cambridge Analytica's illicit procurement of millions of its users' private data to inform data-targeting strategies in the 2016 US presidential election,has shown that there is a crisis in ethics and public acceptance in the data collection companies. Among the many issues raised by that scandal, a subset includes:

- a realization of the massive collection of data beyond the comprehension of the ordinary user;
- the corporate capacity to collate internal and external data and analyze it to achieve personally recognizable data profiles of users, which the users neither knew about nor explicitly approved;
- the collecting of people's data without any contractual or other authority to do so; and
- the lack of transparency in the data collection processes, sources, detail, purposes and use.

These issues are more urgent when they have a direct impact on people's working lives. It is important, to meet the pressing needs of data accuracy and worker confidence, that employees and contractors have access to the data being collected for enterprise AI, and, in particular, for workplace AI. Data quality improves when many eyes have it under scrutiny. Furthermore, to preserve their workplace morale, workers need to be sure that their own personal information is being treated with respect and in accordance with laws on privacy and labour rights.

Including Community Interests
The present discussion about the ethics of data gathering and algorithmic decision making has focused on the rights of individuals. The principles for the adoption of AI need to include an expression of the policy concerns of the community as a whole, as well as those of individuals. For instance, the individual right of intellectual property protection may need to be traded off

_____

[7] Erin Dalton, deputy director of Allegheny County's Department of Human Services, quoted in Hurley (2018).

against the community interest of non-discrimination and, hence, a requirement for greater transparency as to the purpose, as well as the inputs and outputs, of a particular algorithmic decision-making tool.

Risk of Further Marginalization of the Vulnerable

AI, at its heart, is a system of probability analysis for presenting predictions about certain possible outcomes. Whatever the use of different tools for probability analysis, the problem of outliers remains. In a world run by algorithms, the outlier problem has real human costs. A society-level analysis of the impact of big data and AI shows that their tendency toward profiling and limited-proof decisions results in the further marginalization of the poor, the Indigenous and the vulnerable (see Obar and McPhail 2018).

One account reported by Virgina Eubanks (2018, 11) explains how interrelated systems reinforce discrimination and can narrow life opportunities for the poor and the marginalized:

> What I found was stunning. Across the country, poor and working-class people are targeted by new tools of digital poverty management and face life-threatening consequences as a result. Automated eligibility systems discourage them from claiming public resources that they need to survive and thrive. Complex integrated databases collect their most personal information, with few safeguards for privacy or data security, while offering almost nothing in return. Predictive models and algorithms tag them as risky investments and problematic parents. Vast complexes of social service, law enforcement, and neighborhood surveillance make their every move visible and offer up their behavior for government, commercial, and public scrutiny.

This excerpt highlights the issue of unintended consequences, particularly costly when they impact the marginalized. It is unlikely that the code-writers of the systems described above started off with the goal "let's make life more difficult for the poor." However, by not appreciating the power of the outcome of the semi-random integration of systems — each system narrowly incented by the desired outcomes for the common and the privileged — that is exactly what these programmers did.

The same concerns apply to the workplace. As one example, at first glance it may appear intuitive to record how far an applicant lives from the workplace for an algorithm designed to determine more likely long-term employees. But this data inherently discriminates against poorer applicants dependent on cheaper housing and public transport. As another, AI written around a narrow definition of completed output per hour may end up discriminating against slower older employees, whose experience is not reflected in the software model.

Over the past few decades, many employers have adopted corporate social responsibilities, partly in the recognition that their contribution to society is more than just profitability. As the AI revolution continues, it is essential that a concerted effort be made to ensure that broader societal responsibilities are not unwittingly eroded through the invisible operation of narrowly written deterministic algorithms that reinforce each other inside and beyond the enterprise.

Big data and AI should not result in some sort of poorly understood, interlinked algorithmic Benthamism, where the minority is left with diminished life opportunities and further constrained autonomy.

2.5   Humans Are Accountable for AI

There is a tendency by some to view AI, because of its complex and opaque decision making, as being separate from other products made by humans, and a unified entity unto itself. Such a notion is a grave error and one that fails to understand the true role of the human within the algorithm. It is essential to emphasize the human agency within the building, populating and interpretation of the algorithm. Humans need to be held accountable for the product of algorithmic decision making. As Lorena Jaume-Palasí and Matthias Spielkamp (2017, 6-7) state:

> The results of algorithmic processes…are patterns identified by means of induction. They are nothing more than statements of probability. The patterns identified do not themselves constitute a conclusive judgment or an intention. All that patterns do is suggest a particular (human) interpretation and the decisions that follow on logically from that interpretation. It therefore seems inappropriate to speak of "machine agency", of machines as subjects capable of bearing "causal responsibility"...While it is true that preliminary automated decisions can be made by means of algorithmic processes (regarding the ranking of postings that appear on a person's Facebook timeline, for example), these decisions are the result of a combination of the intentions of the various actors who (co-)design the algorithmic processes involved: the designer of the personalization algorithm, the data scientist who trains the algorithm with specific data only and continues to co-design it as it develops further and, not least, the individual toward whom this personalization algorithm is directed and to whom it is adapted. All these actors have an influence on the algorithmic process. Attributing causal responsibility to an automated procedure — even in the case of more complex algorithms — is to fail to appreciate how significant the contextual entanglement is between an algorithm and those who co-shape it.

2.6   A Human-centric Model Is Essential for Acceptance of AI and to Ensure a Safe AI Future

Hundreds of technical and scientific leaders have warned of the risk of integrated networks of AI superseding human controls unless governments intervene to ensure human control is mandated in AI development. The British physicist Stephen Hawking spoke of the importance of regulating AI: "Unless we learn how to prepare for, and avoid, the potential risks, AI could be the worst event in the history of our civilization. It brings dangers, like powerful autonomous weapons, or new ways for the few to oppress the many" (quoted in Clifford 2017); further, he warned, "it would take off on its own, and re-design itself at an ever increasing rate. Humans,

who are limited by slow biological evolution, couldn't compete, and would be superseded" (quoted by Cellan-Jones 2014).

More specifically within the workplace, big data and AI could result in a new caste system imposed on people by systems determining and limiting their opportunities or choices in the name of the code-writers' assumptions about the best outcome for the managerial purpose. One can imagine an AI-controlled recruitment environment where the freedom of the person to radically change careers is punished by algorithms only rewarding commonly accepted traits as being suitable for positions.

AI should not be allowed to diminish the ability of people to exercise autonomy in their working lives and in determining the projection of their own life paths. This autonomy is an essential part of what makes us human. As UNI Global Union (2018, 9) says, in the deployment of these technologies, workplaces should "show respect for human dignity [and] privacy and the protection of personal data should be safeguarded in the processing of personal data for employment purposes, notably to allow for the free development of the employee's personality as well as for possibilities of individual and social relationships in the work place."

Microsoft (2018, 136) has called for a "human-centered approach" to AI. This approach is important not only to control AI's potential power, but to ensure — particularly in the workplace, including the gig economy — that AI serves the values and rights humans have developed as individuals in societies over the last centuries.

As *The Economist* (2018, 13) has concluded: "The march of AI into the workplace calls for trade-offs between privacy and performance. A fairer, more productive workforce is a prize worth having, but not if it shackles and dehumanises employees. Striking a balance will require thought, a willingness for both employers and employees to adapt, and a strong dose of humanity."

2.7   The Need for a Governance Framework

The Facebook crisis has shown how government's role in protecting the rights and well-being of citizens and workers lagged behind the market-driven incentives for companies to conduct large-scale, detailed, unaccountable and shared surveillance of millions of people. The potential disruption of AI signals that it is best, both for business certainty and worker adaption, that this governance lag not be repeated. In an environment where changes to the scope, content, control and reward of work are accelerating, ensuring that workers' apprehensions are addressed in an open and accountable way will be important for ensuring ongoing productivity improvements and avoiding unintended social disruptions. Now is the time for G20 governments to establish a set of principles to guide the adoption of AI and automation in the workplace.

Building on the thinking of companies, think tanks, unions, academics and analytical media,[8] the following set of principles on data collection and AI in the workplace are proposed for consideration by the G20 in Buenos Aires.

3   The Framework Principles

The first set of seven framework principles relates to the collection of data in the work environment.

Right to know data is being collected, for what and from where: Workers, be they employees or contractors, or prospective employees and contractors, must have the right to know what data is being collected on them by their employers, for what purpose and from what sources.

Right to ensure worker data is accurate and compliant with legal rights to privacy: An important feature for worker understanding and productivity is to ensure that workers, ex-workers and job applicants have access to the data held on them in the workplace or have the means to ensure that the data is accurate and can be rectified, blocked or erased if it is inaccurate or breaches legally established rights to privacy. The collection and processing of biometric data and other personally identifiable information (PII) must be proportional to its stated purpose, based on scientifically recognized methods, and held and transmitted very securely.

Principle of proportionality: The data collected on present or prospective employees or contractors should be proportional to its purpose. As one group has proposed: "Collect data and only the right data for the right purposes and only the right purposes, to be used by the right people and only the right people and for the appropriate amount of time and only the appropriate amount of time."

Principle of anonymization: Data should be anonymized where possible. Data with PII should only be available where it is important to the data collection's prime purpose, and its visibility must be limited to the employee and the relevant manager. Aggregated, anonymized data is preferable for many management and productivity purposes.

Right to be informed about the use of data: Employees and contractors should be fully informed when either internal or external data (or both) has been used in a decision affecting their career. Any data processing of present or prospective employees' or contractors' data should be transparent and the PIIavailable for their review. The right to understand and appeal against both the rationale employed and the data used to achieve that rationale is essential to

---

[8] These works are outlined in the bibliography.

safeguard present or prospective workers against poor or inaccurate input data or discriminative decisions.

Limits to monitoring of the workplace by employers: Proportional data collection and processing should not be allowed to develop into broad-scale monitoring of employees or contractors. While monitoring can be an indirect consequence of steps taken to protect production, health and safety or to ensure the efficient running of an organization, continuous general monitoring of workers should not be the primary intent of the deployment of workplace technology. Given the potential in the use of such technology to violate the rights and freedoms of the persons concerned, employers must be actively engaged to ensure that the use is constrained to specific positive purposes, so as not to breach these rights. This principle is not only a matter of workplace freedoms, but also a practical step toward maintaining morale and productivity.

Accuracy of data inputs and the "many eyes" principle: Employers should ensure the accuracy, both in detail and its intended purpose, of the data models and sources for AI. Poor data results in flawed decision making. Training data and training features should be reviewed by many eyes to identify possible flaws and to counter the "garbage in, garbage out" trap. There should be a clear and testable explanation of the type and purpose of the data being sourced. Workers and contractors with experience of the work processes and data environment of the firm should be incorporated into the review of data sources. Such data should be regularly reviewed for accuracy and fit for purpose. Algorithms used by firms to hire, fire and promote should be regularly reviewed for data integrity, bias and unintended consequences.

An additional seven principles focus on AI in the workplace:

Focus on humans: Human control of AI should be mandatory and testable by regulators.

AI should be developed with a focus on the human consequences as well as the economic benefits. A human impact review should be part of the AI development process, and a workplace plan for managing disruption and transitions should be part of the deployment process. Ongoing training in the workplace should be reinforced to help workers adapt. Governments should plan for transition support as jobs disappear or are significantly changed.

Shared benefits: AI should benefit as many people as possible. Access to AI technologies should be open to all countries. The wealth created by AI should benefit workers and society as a whole as well as the innovators.

Fairness and inclusion: AI systems should make the same recommendations for everyone with similar characteristics or qualifications. Employers should be required to test AI in the workplace on a regular basis to ensure that the system is built for purpose and is not harmfully influenced by bias of any kind — gender, race, sexual orientation, age, religion, income, family status and so on. AI should adopt inclusive design efforts to anticipate any potential deployment issues that could unintentionally exclude people. Workplace AI should be tested to ensure that it does not discriminate against vulnerable individuals or communities.

Governments should review the impact of workplace, governmental and social AI on the opportunities and rights of poor people, Indigenous peoples and vulnerable members of society.

In particular, the impact of overlapping AI systems toward profiling and marginalization should be identified and countered.

Reliability: AI should be designed within explicit operational requirements and undergo exhaustive testing to ensure that it responds safely to unanticipated situations and does not evolve in unexpected ways. Human control is essential. People-inclusive processes should be followed when workplaces are considering how and when AI systems are deployed.

Privacy and security: Big data collection and AI must comply with laws that regulate privacy and data collection, use and storage. AI data and algorithms must be protected against theft, and employers or AI providers need to inform employees, customers and partners of any breach of information, in particular PII, as soon as possible.

Transparency: As AI increasingly changes the nature of work, workers, customers and vendors need to have information about how AI systems operate so that they can understand how decisions are made. Their involvement will help to identify potential bias, errors and unintended outcomes. Transparency is not necessarily nor only a question of open-source code. While in some circumstances open-source code will be helpful, what is more important are clear, complete and testable explanations of what the system is doing and why.

Intellectual property, and sometimes even cyber security, is rewarded by a lack of transparency. Innovation generally, including in algorithms, is a value that should be encouraged. How, then, are these competing values to be balanced?

One possibility is to require algorithmic verifiability rather than full algorithmic disclosure. Algorithmic verifiability would require companies to disclose not the actual code driving the algorithm but information allowing the *effect* of their algorithms to be independently assessed. In the absence of transparency regarding their algorithms' purpose and actual effect, it is impossible to ensure that competition, labour, workplace safety, privacy and liability laws are being upheld.[9]

When accidents occur, the AI and related data will need to be transparent and accountable to an accident investigator, so that the process that led to the accident can be understood.

Accountability: People and corporations who design and deploy AI systems must be accountable for how their systems are designed and operated. The development of AI must be responsible, safe and useful. AI must maintain the legal status of tools, and legal persons need to retain control over, and responsibility for, these tools at all times.

Workers, job applicants and ex-workers must also have the "right of explanation" when AI systems are used in human-resource procedures, such as recruitment, promotion or dismissal.[10] They should also be able to appeal decisions by AI and have them reviewed by a human.

_____

[9] This is explored to some degree by the the Global Commission for Internet Governance (2016, 45).

[10] The European Union's General Data Protection Regulation seems to infer a "right to explanation." See Burt (2017).

4   Going Forward

This paper offers principles for G20 governments to consider in enabling the smoother and more socially acceptable introduction of big data and AI into the workplace.

There are two paths toward the adoption of a G20 framework for AI in the workplace.

First, building on the G20 Roadmap for Digitalisation, the ministers responsible for the digital economy could consider the principles outlined in this paper. Think tank participants in the Think 20 Summit engagement group could work with officials to prepare a document for consideration by the second meeting of the Digital Economy Task Force on August 21-22, 2018.

Second, and not inconsistently with the first path, ministers could consider establishing a multi-stakeholder grouping from with the G20 process to flesh out more details of the principles outlined in this paper. This group could report to ministers during the Japanese presidency of the G20 in 2019. Drawing on the expertise of the think tank, business and labour engagement groups of the G20, AI designers and developers, researchers, employers, consumer organizations, lawyers, unions and government officials could work on a more detailed framework for principles, monitoring procedures and compliance process recommendations.

References

Angwin, Julia, Jeff Larson, Surya Mattu and Lauren Kirchner. 2016. "Machine Bias." *ProPublica*, May 23. www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Barghava, Rahul. 2017. "The Algorithms Aren't Biased, We Are." *MIT Media Lab* (blog), January 3. https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2.

Burt, Andrew. 2017. "Is There a 'Right to Explanation' for Machine Learning in the GDPR?" IAPP *Privacy Tech* (blog), June 2. International Association of Privacy Professionals. https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/

Cellan-Jones, Rory. 2014. "Stephen Hawking Warns Artificial Intelligence Could End Mankind." BBC News, December 2. www.bbc.com/news/technology-30290540

Clifford, Catherine. 2017. "Hundreds of A.I. Experts Echo Elon Mush, Stephen Hawking in Call for a Ban on Killer Robots." CNBC, November 8. https://www.cnbc.com/2017/11/08/ai-experts-join-elon-musk-stephen-hawking-call-for-killer-robot-ban.html

The Economist. 2018. "AI-Spy: The Workplace of the Future." March 28. https://www.economist.com/leaders/2018/03/28/the-workplace-of-the-future

Eubanks, Virgina. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* New York, NY: St. Martin's Press.

Frey, Carl Benedikt and Michael A. Osborne. 2013. *The Future of Employment: How Susceptible Are Jobs to Computerisation?* September 17. Oxford, UK: Oxford Martin Programme on Technology and Employment. www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

G20. 2017a. *G20 Digital Economy Ministerial Conference. Düsseldorf 6-7 April 2017.* Declaration of the Ministers Responsible for the Digital Economy. Federal Ministry for Economic Affairs and Energy. www.bmwi.de/Redaktion/DE/Downloads/G/g20-digital-economy-ministerial-declaration-english-version.pdf?__blob=publicationFile&v=12

G20. 2017b. "A ROADMAP for Digitalisation: Policies for a Digital Future." Annex paper 1 to the Declaration of the Ministers responsible for the Digital Economy. In *G20 Digital Economy Ministerial Conference. Düsseldorf 6-7 April 2017*, 10–15. www.bmwi.de/Redaktion/DE/Downloads/G/g20-digital-economy-ministerial-declaration-english-version.pdf?__blob=publicationFile&v=12

G20. 2017c. "G20 Leaders' Declaration: Shaping an Interconnected World." G20 Germany 2017 meetings, Hamburg, July 7-8. www.g20germany.de/Content/EN/_Anlagen/G20/G20-leaders-declaration.pdf;jsessionid=0C08AA235271BF43ECBB08BA059EE5B7.s6t2?__blob=publicationFile&v=11

Gangadharan, Seeta P., Virginia Eubanks and Solon Barocas, eds. 2014. *Data and Discrimination: Collected Essays*. Washington, DC: New America. www.newamerica.org/oti/policy-papers/data-and-discrimination/

Global Commission on Internet Governance. 2016. *One Internet: Final Report of the Global Commission on Internet Governance.* Waterloo, ON: CIGI. www.cigionline.org/publications/one-internet

Hurley, Dan. 2018. "Can an Algorithm Tell When Kids Are in Danger?" *New York Times*, January 2. www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html

Jaume-Palasí, Lorena and Matthias Spielkamp. 2017. "Ethics and Algorithmic Processes for Decision Making and Decision Support." AlgorithmWatch Working Paper No. 2, June 1. Berlin, Germany: AlgorithmWatch. https://algorithmwatch.org/en/ethics-and-algorithmic-processes-for-decision-making-and-decision-support/

KPMG International. 2016. "Rise of the Humans: The Integration of Digital and Human Labor." KPMG International Cooperative, November. https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2016/11/rise-of-the-humans.pdf

Lohr, Steve. 2017. "A.I. Will Transform the Economy. But How Much, and How Soon?" *New York Times*, November 30. www.nytimes.com/2017/11/30/technology/ai-will-transform-the-economy-but-how-much-and-how-soon.html

Manyika, James, Susan Lund, Michael Chui, Jacques Bughin, Jonathan Woetzel, Parul Batra, Ryan Ko and Saurabh Sanghvi. 2017. *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation.* San Francisco, CA: McKinsey Global Institute. www.mckinsey.com/featured-insights/future-of-organizations-and-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages

Microsoft. 2018. *The Future Computed: Artificial Intelligence and Its Role in Society.* Redmond, WA: Microsoft. https://blogs.microsoft.com/uploads/2018/02/The-Future-Computed_2.8.18.pdf

Obar, Jonathan, and Brenda McPhail. 2018. "Preventing Big Data Discrimination in Canada: Addressing Design, Consent and Sovereignty Challenges." In *Data Governance in the Digital Age*: *Special Report*, 56–64. Waterloo, ON: CIGI. https://www.cigionline.org/publications/data-governance-digital-age.

Ovenden, James. 2016. "AI In Developing Countries: Artificial Intelligence Isn't just for Self Driving Cars." Innovation Enterprise, October 6. https://channels.theinnovationenterprise.com/articles/ai-in-developing-countries.

Purdy, Mark, and Paul Daugherty. 2016. "Why Artificial Intelligence Is the Future of Growth." Accenture Institute for High Performance, September 28. www.accenture.com/us-en/insight-artificial-intelligence-future-growth

UNI Global Union. 2018. *Top 10 Principles for Workers' Data Privacy and Protection.* Nyon, Switzerland. www.thefutureworldofwork.org/docs/10-principles-for-workers-data-rights-and-privacy/

Wexler, Rebecca. 2017. "When a Computer Program Keeps You in Jail." *New York Times*, June 13. www.nytimes.com/2017/06/13/opinion/how-computers-are-harming-criminal-justice.html

Please note:

You are most sincerely encouraged to participate in the open assessment of this discussion paper. You can do so by either recommending the paper or by posting your comments.

Please go to:

The Editor