FUTURE OF WORK AND EDUCATION FOR THE DIGITAL AGE

Building on the Hamburg Statement and The G20 Roadmap For Digitalization – Towards a G20 Framework For Artificial Intelligence in the Workplace

Dr Paul Twomey, Centre for International Governance Innovation (CIGI)

T20
ARGENTINA 2018
THINK 20

www.t20argentina.org

/T20Solutions

@T20Solutions

/T20Solutions

CARI / CONSEJO ARGENTINO PARA LAS
RELACIONES INTERNACIONALES

CIPPEC

Abstract

Building on the 2017 Hamburg Statement and the G20 Roadmap for Digitalization, this Policy Brief 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.

Challenge

In their 2017 Hamburg Statement[1], G-20 leaders 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."

Leaders also recognised that the swift adoption of ICT 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."

Responding to the rise of Big Data and Artificial Intelligence (AI)[2] is one of the most important ways that G-20 leaders could address their Hamburg Statement goals.

---

1 https://www.g20germany.de/Content/EN/_Anlagen/G20/G20-leaders-declaration.html;jsessionid=0C16852EEB7D1B06ADE151F7FDCE44FB.s6t2?nn=2186554

2 Nearly all software contains some form of algorithm and poses little disruption to the workplace. But the complex algorithms that drive significant decision making in the workplace have drawn public attention. For the purpose of this policy brief I will use the term Artificial Intelligence (AI) to cover automated decision-making informed by complex algorithms and machine learning capabilities

**The Future of Work
and Education
for the Digital Age**

T20
ARGENTINA 2018
THINK 20

In Hamburg, leaders stated that the "G20 Roadmap for Digitalisation will help us guide our future work." In that Roadmap, Ministers responsible for the Digital Economy, said that they would further discuss "frameworks as enablers for... workforce digitalization." Some aspects for such frameworks were indicated: "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 Argentinian Presidency of the G20 we will discuss international public policy issues related to privacy and security in the digital economy."[3]

These issues –trust, security, the need to adapt, privacy, skills– all are central as workers and citizens react to the rapid introduction of Big Data collection and related AI. Confronted with forecasts that these technologies may affect up to nearly half of all jobs[4], workers worry about their employment and what skills they will need. People seek assurance that the introduction of AI and automation will be in a manner which ensures respect for the human integrity of workers and is done under a framework of accountability while still delivering the productivity, safety, and innovation benefits promised.

This Policy Brief offers such a framework for G-20 governments to enable the smoother and more socially acceptable introduction of Big Data and AI. It explores the main issues and proposes framework principles. It also suggests two paths towards adoption of a G-20 Framework for Artificial Intelligence in the Workplace.

A detailed discussion of the issues which are addressed by the principles in this framework is attached in the Appendix.

---

3   https://www.bmwi.de/Redaktion/DE/Downloads/G/g20-digital-economy-ministerial-declaration-english-version.pdf?__blob=publicationFile&v=12

4   KPMG reports: 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 UK economy in the next 20 years. Digital technologies will conceivably off-set the jobs of 130 million knowledge workers — or 47 per cent of total US employment — by 2025. Across the OECD some 57 per cent of jobs are threatened. In China, that number soars to 77 per cent. KPMG, The Rise of the Humans, 2016 https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2016/11/rise-of-the-humans.pdf

**The Future of Work
and Education
for the Digital Age**

T20
ARGENTINA 2018
THINK 20

Proposal

Building on the thinking of companies, think tanks, unions, academics and analytical media[5], the following set of principles on data collection and AI in the workplace are proposed for consideration by the G-20 in Buenos Aires.

The Framework Principles

Data collection 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 that workers, ex-workers and job applicants have access to the data held on them in the workplace and/or have 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 Identifying Information (PII) must be proportional to its stated purpose, based on scientifically recognised methods and be held and transmitted very securely.

Principal 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."

Principal of Anonymization
Data should be anonymized where possible. Personally Identifying Data should only be available where it is important to the data collection's prime purpose, and its visibility limited only 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 internal and/or external

---

5 These works are outlined in the bibliography.

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

data has been used in a decision effecting their career. Any data processing of present or prospective employees' or contractors' data should be transparent and the Personally Identifying Information be available 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 safeguard present or prospective workers against poor/inaccurate input data and/or discriminative decisions.

Monitoring of the workplace by employers should be limited to specific positive purposes

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 to violate the rights and freedoms of persons concerned by the use of such technology, employers must be actively engaged to ensure that the use is constrained so as not to breach these rights. This in not just a matter of workplace freedoms, it is also a practical step towards 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. The establishment of 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 re 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.

Artificial Intelligence in the workplace

Human-focused

Human control of AI should be mandatory and be 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.

Benefits should be shared

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 - be it gender, race, sexual orientation, age, religion, income, family status etc. AI should adopt inclusive design efforts to anticipate any potential deployment issues that could unintentially exclude people. Workplace AI should be tested to ensure that it does not discriminate against the vulnerable in our communities.

Governments should review the impact of workplace, governmental and social AI on the opportunities and rights of the poor, the indigenous and the vulnerable. In particular the impact of overlapping AI systems towards 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, 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, especially Personally Identifying Information as soon as possible.

Transparency

As AI increasingly changes the nature of work, workers and customers/vendors need to have information about how AI systems operate so that they can understand how decisions are made. Their involvement will help identify potential bias, errors and unintended outcomes. Transparency is not necessarily nor sufficiently just a

question of open source code. While in some circumstances this will be helpful, what is more important is clear, complete and testable explanations of what the system is doing and why.

Intellectual property and sometimes even cybersecurity is rewarded by a lack of transparency. Innovation generally, including in algorithms, is a value which should be encouraged. How 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 information allowing the effect of their algorithms to be independently assessed, but not the actual code driving the algorithm. Without such transparency as to purpose and actual effect it is impossible to ensure that competition, labour, workplace safety, privacy and liability laws are being upheld.[6]

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[7]. They should also be able to appeal decisions by AI and have them reviewed by a human.

Going Forward

This Policy Brief offers principles for G-20 governments to consider in enabling the smoother and more socially acceptable introduction of Big Data and AI into the workplace.

---

6 This is explored in some degree by the Report of the Global Commission for Internet Governance, p 45 https://www.ourinternet.org/report

7 The EU General Data Protection Regulation (GDPR) seems to infer a "right to explanation". See Andrew Burt, "Is there a 'right to explanation' for machine learning in the GDPR?, International Association of Privacy Professionals, 1 June 2017 https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/

**The Future of Work
and Education
for the Digital Age**

T20

ARGENTINA 2018
THINK 20

There are two paths towards the adoption of a G-20 Framework for AI in the Workplace.

First, building on the G20 Roadmap for Digitalisation Ministers responsible for the Digital Economy could consider the principles outlined in this Policy Brief. T-20 participants could work with officials to prepare a document for consideration by the 2nd Meeting of the Digital Economy Task Force on 21-22 August 2018.

Secondly, and not inconsistently with the first path, Ministers could consider establishing a multi-stakeholder grouping from with the G-20 process to flesh out more details of the principles outlined in this Policy Brief. This group could report to Ministers as part of the Japanese presidency of the G-20. Drawing on the T20, B20, and L20, AI designers and developers, researchers, employers, consumer organisations, lawyers, unions and government officials could work on a more detailed framework for principles, monitoring procedures and compliance process recommendations.

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

References

1. Allesando Acquisti and Christina M. Fong. "An experiment in hiring discrimination via online social networks." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2031979. 2015

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

3. Rahul Barghava, "The algorithms are not biased, we are", MIT Media Lab, January 3, 2017 https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2

4. Solon Barocas and Andrew D. Selbst. "Big Data's disparate impact." California Law Review 104: 671–732.

5. British Columbia First Nations Data Governance Initiative. Decolonizing Data: Indigenous Data Sovereignty Primer, April 2017.

6. Andrew Burt, "Is there a 'right to explanation' for machine learning in the GDPR?" International Association of Privacy Professionals, 1 June 2017 https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/

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

8. Rory Elnnan-Jone, "Stephen Hawking warns artificial intelligence could end mankind", BBC News, 2 December 2014 http://www.bbc.com/news/technology-30290540

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

10. Carl Benedikt Frey and Michael A. Osborne, The Future of Employment: How Susceptible Are Jobs to Computerisation?, Published by the Oxford Martin Programme on Technology and Employment, 2013

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

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

12. Global Commission on Internet Governance, One Internet: Final Report of the Global Commission on Internet Governance, (London, Centre for International Governance Innovation and The Royal Institute for International Affairs, 2016) https://www.ourinternet.org/report

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

14. Lorena Jaume-Palasí and Matthias Spielkamp, "Ethics and algorithmic processes for decision making and decision support", AlgorithmWatch Working Paper No. 2 pp 6-7. https://algorithmwatch.org/en/ethics-and-algorithmic-processes-for-decision-making-and-decision-support/

15. Lauren Kirchner, "New York City moves to create accountability for algorithms." Ars Technica, December 19, 2107. https://arstechnica.com/tech-policy/2017/12/new-york-city-moves-to-create-accountability-for-algorithms/.

16. KPMG, The Rise of the Humans, 2016 https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2016/11/rise-of-the-humans.pdf

17. Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson and Harlan Yu, "Accountable algorithms." University of Pennsylvania Law Review 165, 2017 pp.633–705.

18. Seve Lohr, "A.I. will transform the Economy. But How Much, and How Soon/", New York Times, November 30, 2017.

19. Mary Madden, Michele Gilman, Karen Levy and Alice Marwick, "Privacy, poverty, and Big Data: A matrix of vulnerabilities for poor Americans." Washington University Law Review 95 pp.53–125, 2017.

20. McKinsey Global Institute, "Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation", 2017 https://www.mckinsey.com/~/media/McKinsey/Global%20Themes/Future%20of%20Organizations/What%20the%20future%20of%20work%20will%20mean%20for%20jobs%20skills%20and%20wages/MGI-Jobs-Lost-Jobs-Gained-Report-December-6-2017.ashx

21. Microsoft, The Future Computed, Redmond, 2017, https://msblob.blob.core.windows.net/ncmedia/2018/02/The-Future-Computed_2.8.18.pdf

22. Safiya U. Noble, Algorithms of Oppression: How Search Engines Reinforce Racism.

**The Future of Work
and Education
for the Digital Age**

T20
ARGENTINA 2018
THINK 20

New York, NY: New York University Press, 2018.

23. Jonathan Obar and Brenda McPhail, Preventing Big Data Discrimination in Canada: Addressing Design, Consent and Sovereignty Challenges, (Wellington, Centre for International Governance Innovation, 2018) https://www.cigionline.org/articles/preventing-big-data-discrimination-canada-addressing-design-consent-and-sovereignty

24. Jonathan A. Obar and Anne Oeldorf-Hirsch, "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services." 2016 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2757465. Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, NY: Broadway Books, 2017.

25. James Ovendon, "AI in Developing Countries: Artificial Intelligence is not just for driverless cars", Innovation Enterprise, 6 October 2016, https://channels.theinnovationenterprise.com/articles/ai-in-developing-countries

26. Frank Pasquale, The Black Box Society: The Secret Algorithms that Control Money and Information. Cambridge, MA: Harvard University Press, 2015.

27. Mark Purdy and Paul Daugherty, Why Artificial Intelligence is the Future of Growth, Accenture, https://www.accenture.com/us-en/insight-artificial-intelligence-future-growth

28. Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, Rohan Ramanath, N. Cameron Russell, Norman Sadeh and Florian Schaub, "Disagreeable privacy policies: Mismatches between meaning and users' understanding." Berkeley Technology Law Journal 30 (1) 2015, pp 39–68.

29. Christian Sandvig, Kevin Hamilton, Karrie Karahalios and Cedric Langbort, "When the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software." International Journal of Communication 10, 2016, pp. 4972–90.

30. R. Joshua Scannell, "Broken windows, broken code." Reallifemag.com, August 29, 2016 http://reallifemag.com/broken-windows-broken-code/.

31. Daniel J. Solove, "Introduction: Privacy self-management and the consent dilemma." Harvard Law Review 126, 2012, pp. 1880–1903.

32. The White House, Big Data: Seizing Opportunities, Preserving Values. https://

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

33. The Economist, March 31st 2018, Special Edition on Artificial Intelligence

34. Joseph Turow, The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth. New Haven, CT: Yale University Press, 2012.

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

36. UNI Global Union, Top 10 Principles for Ethical Artificial Intelligence, Nyon, Switzerland , 2017 http://www.thefutureworldofwork.org/opinions/10-principles-for-ethical-ai/

The Future of Work
and Education
for the Digital Age

T2O
ARGENTINA 2018
THINK 20

Appendix

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 the population, algorithms are determining our present rights and future opportunities. To take just land transport, they help drive our cars, determine whether we can get a loan to buy our cars, decide which roads should be repaired, identify if we have broken the road rules, even determine whether we should be imprisoned if we have[8].

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 unclear outcomes than humans making decisions. Examples include the prevention of medical errors to increasing productivity and reducing risks in the workplace. Machine learning can improve job descriptions and provide more "blind" recruitment processes which can increase both the pool of qualified candidates and boost recruitment of non-conventional applicants[9]. Written well, algorithms can be more impartial and pick up patterns people may miss.

Many commentators point to the productivity benefits of AI. For instance, analysis by Accenture of twelve 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 productivity by up to 40 percent and enable people to make more efficient use of their time."[10] The World Bank is exploring the benefits of AI for development.[11] Others identify farming, resource provision and healthcare as sectors in the developing economies which will benefit greatly from the application of AI.[12]

---

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

9 See firms like Textio and Pymetrics

10 Mark Purdy and Paul Daugherty, Why Artificial Intelligence is the Future of Growth, Accenture, https://www.accenture.com/us-en/insight-artificial-intelligence-future-growth

11 See https://www.measuredev.org/

12 See James Ovendon, "AI in Developing Countries: Artificial Intelligence is not just for driverless cars", Innovation Enterprise, 6 October 2016, https://channels.theinnovationenterprise.com/articles/ai-in-developing-countries

Impact on employment

Much has been made of the impact of AI and related robotics on jobs, especially since Osborne and Frey's 2013 article estimating that 47 percent of jobs in the US were "at risk" of being automated in the next 20 years.[13] Debate has ensured on the exact nature of this impact: the full or part erosion of existing job tasks, the impact across sectors and across developed, emerging and developing economies. Forecasting such things is inherently difficult. But a recent summary by the McKinsey Global Institute reflects a mid-way 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.[14]

Whatever the specifics, the result is clearly going to be very significant for G-20 economies and their citizens. And if the pace of adoption continues to outpace

---

13 Carl Benedikt Frey and Michael A. Osborne, The Future of Employment: How Susceptible Are Jobs to Computerisation?, Published by the Oxford Martin Programme on Technology and Employment, 2013

14 McKinsey Global Institute, "Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation", 2017 https://www.mckinsey.com/~/media/McKinsey/Global%20Themes/Future%20of%20Organizations/What%20the%20future%20of%20work%20will%20mean%20for%20jobs%20skills%20and%20wages/MGI-Jobs-Lost-Jobs-Gained-Report-December-6-2017.ashx

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

previous major technological adoption[15], the scale of social dislocation is likely to be greater. Even more reason for the G-20 to work now on a framework for AI adoption.

The risk of bias

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

As Airbnb says[16], bias in the writing of algorithms is inevitable. It can have chilling effects on individual rights, choices, and the application of worker and consumer protections. 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. Algorithms are not neutral: they incorporate built-in values and serve business models that may lead to unintended biases, discrimination or economic harm[17]. 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, and they may not even be aware of this power, or the potential harm that an incorrectly coded algorithm may have. Because the complex market of interacting algorithms continues to evolve, it's also likely that existing algorithms that may have been innocuous yesterday will have significant impact tomorrow.

AI can present two big flaws:
- bias in its coding, or
- selection bias in or distortion/corruption in its data inputs.

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

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 has a dominant impact on the efficacy of programming. In

---

15 See discussion in Seve Lohr, "A.I. will transform the Economy. But How Much, and How Soon/", New York Times, November 30, 2017.

16 https://airbnb.design/anotherlens/

17 For instance, media reports have pointed out clear racial bias resulting from reliance on sentencing algorithms used by many US courts. https://www.nytimes.com/2017/06/13/opinion/how-computers-are-harming-criminal-justice.html

simple terms, "garbage in, garbage out". This 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 judgements about what data sets tell us which is not necessarily the case. The establishment of training data and training features is at the heart of AI. As Rahul Barghava says, **"In machine learning, the questions that matter are "what is the textbook" and "who is the teacher."**[18] The more scrutiny these can receive the more likely that the data will be fit for purpose. Some local governments in the US 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 purpose, process and data. And 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."[19] Sometimes the data is just flawed. But the more scrutiny it receives the better that 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 Facebook crisis has shown that there is a crisis in ethics and public acceptance in the data collection companies. Only a subset of issues raised include:
- A realization of the massive collection of data beyond the comprehension of the ordinary user
- Corporate capacity to collate internal and external data and analyse that data to achieve personally recognizable data profile outcomes of which the users did not understand nor explicitly approve
- The collection of data on people with whom there is no contractual or other authority to do so
- Lack of transparency in the data collection processes, sources, detail, purposes, and use

These issues are more pressing when they have a direct impact on people's working lives. It is important for the pressing needs of data accuracy and worker confidence

---

18 Rahul Barghava, "The algorithms are not biased, we are", MIT Media Lab, January 3, 2017 https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2
19 Erin Dalton, Deputy Director of Allegheny County's Department of Human Services quoted in Dan Hurley, "Can an Algorithm Tell When Kids Are in Danger", New York Times, 2 January 2018, https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

that employees and contractors have access to the data being collected for enterprise, especially workplace, AI. Data quality improves when many eyes have it under scrutiny. Furthermore, to preserve workplace morale workers need to know that their own personal information is being treated with respect and in accordance with laws on privacy and labour rights.

Community Interests are not just Individual or Corporate Rights

The present discussion about the ethics of data gathering and algorithmic decision-making has focused on the rights of individuals. The principals for the adoption of AI need also to include an expression of the policy concerns of the community as a whole as well as those of individuals. Four instance the individual right of intellectual property protection may need to be traded off against the community interest in non-discrimination and hence a requirement for greater transparency as to the purpose, 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 its tendency towards profiling and limited proof decisions results in the further marginalisation of the poor, the indigenous and the vulnerable.[20]

One account reported by Virgina Eubanks explains how interrelated systems reinforce discrimination and narrow life opportunities for the poor and marginalised:

> 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

---

20 See Jonathan Obar and Brenda McPhail, Preventing Big Data Discrimination in Canada: Addressing Design, Consent and Sovereignty Challenges, (Wellington, Centre for International Governance Innovation, 2018) https://www.cigionline.org/articles/preventing-big-data-discrimination-canada-addressing-design-consent-and-sovereignty

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

service, law enforcement, and neighborhood surveillance make their every move visible and offer up their behavior for government, commercial, and public scrutiny.[21]

This highlights the issue of unintended consequences, especially 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". But by not appreciating how powerful would be the outcome of the semi-random integration of systems, each 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 work place. At first glance it may be appear intuitive to record how far an applicant lives from the workplace for an algorithm designed to determine more likely long-term employees. But it inherently discriminates against poorer applicants dependent on cheaper housing and public transport. 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 last decades, many employers have adopted Corporate Social Responsibilities partly in recognition that their contribution to society is more than just profitability. It is essential as the AI revolution continues, that a concerted effort is made to ensure that the broader societal responsibilities are not unwittingly eroded through the invisible operation of narrowly written deterministic algorithms which 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.

AI is the ongoing product of humans, and hence they are accountable

With complex and opaque decision making, there is tendency by some to see the AI as being separate and unified entity unto itself. This is a grave error and fails to understand the true role of the human within the algorithm. It essential to emphasise 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.

---

21 Virginia Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor, (New York, St Martin's Press, 2018) p.11

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

As Lorena Jaume-Palasí and Matthias Spielkamp, 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.[22]

A human centric model is essential for acceptance 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 warned of the importance of regulating artificial intelligence: "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,"[23]

"It would take off on its own, and re-design itself at an ever increasing rate. Humans, who

---

22 Lorena Jaume-Palasí and Matthias Spielkamp, "Ethics and algorithmic processes for decision making and decision support", AlgorithmWatch Working Paper No. 2 pp 6-7. https://algorithmwatch.org/en/ethics-and-algorithmic-processes-for-decision-making-and-decision-support/

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

The Future of Work
and Education
for the Digital Age

T20
ARGENTINA 2018
THINK 20

are limited by slow biological evolution, couldn't compete, and would be superseded."[24]

More specifically within the workplace, Big Data and AI could result in a new caste system imposed on people by systems which determine and limit 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 career 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 path. This is an essential part of what makes us human. As UNI Global Union says, in the deployment of these technologies, workplaces should "show respect for human dignity, 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."[25]

Microsoft has called "to put humans at the center" of AI[26]. This is important both to control its potential power, and especially in the work place, including the gig economy, to ensure that AI serves the values and rights we have developed as individuals in societies over the last centuries.

As the Economist 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".[27]

The need for a governance framework

The Facebook crisis has shown how government's role in protecting the rights and wellbeing of citizens and workers lagged behind the solely market driven incentives for companies to conduct large scale, detailed, poorly accountable and shared surveillance of millions of people. The potential disruption of AI signals that it is

---

24 Rory Elnnan-Jone, "Stephen Hawking warns artificial intelligence could end mankind", BBC News, 2 December 2014 http://www.bbc.com/news/technology-30290540

25 Top 10 Principles for Workers' Data Privacy and Protection, UNI Global Union, Nyon, Switzerland , 2018

26 The Future Computed, Redmond, Microsoft, 2017, https://msblob.blob.core.windows.net/ncme-dia/2018/02/The-Future-Computed_2.8.18.pdf

27 "AI-spy The workplace of the future", The Economist, March 31st 2018, p13.

**The Future of Work
and Education
for the Digital Age**

T20
ARGENTINA 2018
THINK 20

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 G-20 governments to establish a set of principles which should guide the adoption of artificial information and automation in the workplace.