ORF ISSUE BRIEF
SEPTEMBER 2019
ISSUE NO. 313

Formulating AI Norms: Intelligent Systems and Human Values

Trisha Ray

ABSTRACT In recent years, various governments, international organisations, civil society groups and technology companies have issued documents outlining their principles around the development and use of Artificial Intelligence (AI). Yet, the world appears to be no closer to a universal set of AI norms. This brief suggests a rethinking of how AI norms should be formulated and outlines key lessons. First, technology firms reflect certain human biases that do not do justice to their global consumer base and make them unsuitable to lead the setting of AI principles. Second, while norms are ambiguous by design, the misuse of this ambiguity by actors to justify rights violations sets a dangerous precedent. Third, no single regulation can account for the consequences of the same AI application deployed in different contexts. Finally, algorithms may need to carry certain biases to alleviate social inequity.

Attribution: Trisha Ray, "Formulating AI Norms: Intelligent Systems and Human Values", *ORF Issue Brief No. 313*, September 2019, Observer Research Foundation.

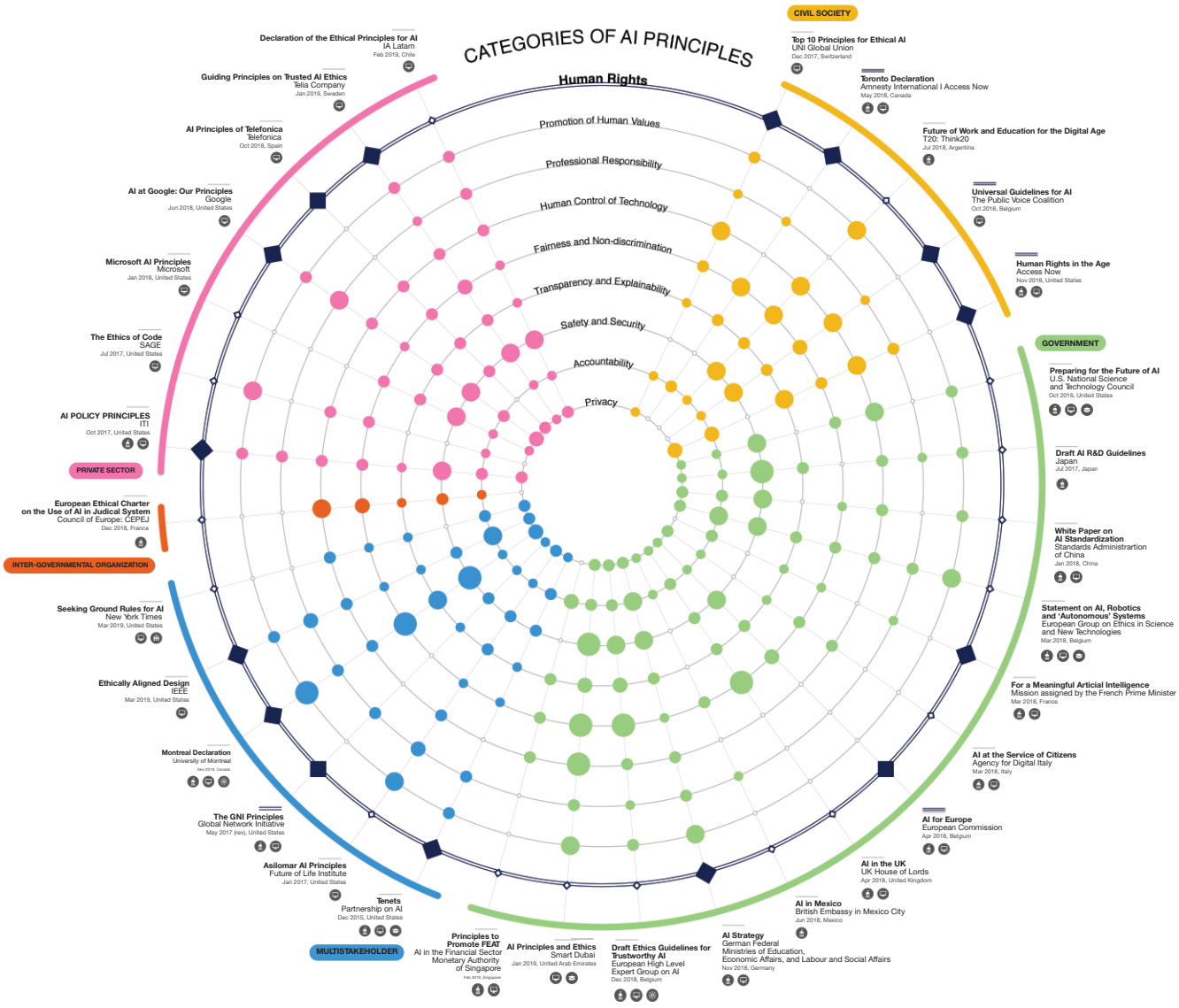
ISBN 978-93-89094-80-0

INTRODUCTION

In the United States (US) in the Fall of 2017, analysts were singing dirges for the country's erstwhile, uncontested technological supremacy, including in the field of Artificial Intelligence (AI). The release of China's ambitious AI strategy, along with data outlining how countries like Japan, South Korea, China, Germany and India were racing to build a body of research, talent and investment had made one analyst warn, "A lot of people take it for granted that the US builds the best tech in the world, and I think that's a dangerous assumption to make."[1] Another American AI expert described his country's conundrum thus: "Shouldn't we take steps to at least slow down progress on AI in the interest of caution? The problem is that if we do so, then nations like China will overtake us."[2]

A couple of years later, the discourse on AI has come a long way. Today there is consensus around the essential principles that must guide the development and use of AI: ethics, transparency, fairness and explainability. Various governments, international organisations, members of civil society as well as technology companies have been issuing documents outlining their principles around AI. (See Figure 1)

Figure 1. Visualising AI Principles



Source: Berkman Klein Center for Internet and Society, Harvard University

Even as there is apparent harmony in these principles, the world is yet to come close to having a set of universal norms for the development and use of AI. This is because underneath these seemingly complementary principles, the contest is intensifying. The creation of AI norms has become a race in itself, and stakeholders are jostling to create a first-mover's advantage by shaping the future trajectory of the technology in ways that secure their own interests. These actors' motivations may include rebuilding or maintaining consumer trust, controlling and in some cases weaponising a powerful tool, or ensuring that certain groups are not left behind.

To be sure, debates around AI norms are a crystallisation of various trends, including growing social and political fissures. They highlight the lack of preparedness of societies for the ubiquity of emerging technologies. There are fundamental questions that need to be answered with regards to the setting of AI norms: What are the constituencies for these norms and do current processes reflect all their interests? Is the creation of robust, universal norms feasible in today's geopolitical and social climate? It is therefore worthwhile to examine developments in the formulation of AI norms over the past couple of years and identify the weaknesses of the process. This brief highlights the failings of AI norms and urges norm entrepreneurs and leaders to go back to the drawing board.

AI NORMS: DEFINITIONS

Existing frameworks around AI norms—whether they be from governments, intergovernmental organisations, private firms or NGOs—broadly fall under one or more of three categories:[3]

a. *Legality:* The design, development and deployment of AI should adhere to any applicable laws and regulations. Legality by definition is a set of enforceable parameters and violations can be punished through clear state-sanctioned mechanisms.

b. *Ethicality:* AI should seek to benefit all of humanity and should not result in any undue harm. Values such as diversity, fairness and social benefit feed into ethicality. Ethics are internally-driven, voluntary practices. Thus, adherence to privacy laws is a lawful action, rather than an ethical one, although the source of a law can of course be an ethical principle.

c. *Robustness:* There should be mechanisms integrated into AI that facilitate all of the above. Ensuring robustness includes technical measures such as explainability, safety and failsafes integrated into the development of algorithms and that assist engineers and programmers in understanding AI behaviour.

A common misconception in discussions about AI norms is that they are "like putting the foundations in before you build a house…now is the time to do it."[4] In reality, the house is well on its way to being built; norms frameworks are like the zoning regulations that will belatedly be put in place. Indeed, AI has already triggered changes in societies and institutions, whether in the form of algorithms that are misused to amplify misinformation and alter the course of an election, as was the case leading up to the 2016 US election, or social credit scores that create a

new hierarchy of citizenship, as with China's Social Credit System.[5] Like zoning regulations that seek to harmonise the multitude of ways that land can be used, norms can only work to ensure that these technologies are used in a way that limits harm and balances competing interests. What have the past couple of years shown about the quest for AI norms?

LESSONS FROM CURRENT AI NORMS-SETTING

1. Self-policing by technology giants is easier said than done.

Technology giants have attempted to take the baton of norm entrepreneur, but are ill-equipped, in part due to loss of consumer trust and also because they reflect a certain set of human biases that fail to do justice to their diverse users.

In March 2019, Google announced the formation of the Advanced Technology External Advisory Council (ATEAC) tasked with "responsible development of AI" for the corporation.[6] Yet, from its inception, the council faced an uphill battle: its eight members, between them, hailed from only four countries, yet were expected to consult on ethical AI development for all of Google's massive global base.

A few days later, one of the eight announced his resignation on Twitter, alongside a veiled condemnation of Google's unpalatable past behavior, which included the (now-scrapped) Project Dragonfly, a censored version of Google for China, and a litany of privacy violations.[7] Concurrently, 3,000 Google employees signed a petition calling for the removal of Kay Cole James, president of the *The Heritage Foundation*, from the council for her organisation's anti-immigrant, anti-LGBTQ+ views. The petition was also signed by other individuals from academia, the tech industry and civil society.[8] Only a little over a week after its announcement, the Council was dissolved. "It's become clear that in the current environment, ATEAC can't function as we wanted," a Google spokesperson said. "We're ending the council and going back to the drawing board."[9]

The episode with Google's AI Ethics Council provides some pertinent lessons. First, while technology platforms like Google have a stake in outlining ethics and norms in order to build trust amongst their users, they also face a Catch-22: the trust deficit they are attempting to close has itself become a barrier to engagement with other key stakeholders.

User trust in platforms has deteriorated in the wake of successive data breaches and scandals. A 2017 Politico survey of registered US voters found the trust ratio of Facebook users at a low 10 percent; that of Google was 13 percent; and for Twitter, eight percent.[10] The chasm is not restricted to the US: a 2018 survey in India found that 56.6 percent of respondents do not trust social media platforms to safeguard their data.[11]

Even as these platforms scramble to repair their public image through renewed commitments to ethics and transparent algorithms, the damage seems to have already been done: lawmakers in many parts of the world are calling for greater oversight of tech giants; some are deploying antitrust measures, and others are calling for data localisation.[12]

A second lesson highlighted by the case of the Google AI Ethics Council is that multinational technology giants like Google are increasingly facing pushback from their own employees. These internal critiques, however, often dismiss rather than address the most important points of contention.

The social biases that may be inherent in tech giants like Google were brought to the fore by the incident involving former Google engineer, James Damore.[13] Damore criticised what he calls "Google's ideological [code word for liberal] echo chamber." He pointed to Google's emphasis on diversity and gender parity as being discriminatory, as they assume men and women have the same capacities. Damore's letter incorrectly conflates equity and discrimination – the many flaws in his argument are well-documented – but it does demonstrate the fairly widespread belief that attempts to rectify historical inequity in the tech sector are threatening to the incumbent, in this case the homogenised category of "conservative white men".

The outcry over Damore's letter and the AI Ethics Council – both within and outside Google – as well as the outrage about the outcry are both a symptom and a cause of intensifying polarisation. While many took to social media to criticise the infamous letter, Damore became an overnight celebrity in alt-right and alt-light circles and his story was used as proof that Silicon Valley was anti-conservative.[14,#] Google's attempt to change this perception by inviting the president of a conservative think tank to its AI Ethics Council was met with backlash from its employees, helping the alt-right to further cement their 'anti-conservative' rhetoric.

These cases demonstrate another problem with AI principle-making as it exists today: the focus on outcomes rather than processes. The outcome focus is ineffective for two reasons: First, social psychologists say that human beings are incapable of thinking about the long-term consequences of their actions.[15] Second, human beings are limited by 'bounded ethicality': people make decisions within the limitations of their own contexts.[16] Thus, Facebook's PR and communications head, Elliot Schrage argued that then-candidate Donald Trump's divisive language did not violate Facebook's terms of use; Facebook public policy head Joel Kaplan stated that Russian meddling would expose Facebook to accusations of being pro-Democrat; and these myopic decisions built up.[17] Blame can be attributed to advertising algorithms for tilting the outcomes of the 2016 US elections, but when Facebook introduced user data-based ad targeting in 2009, no one could have predicted that the tool would be used to influence elections in the world's oldest democracy.[18]

Technology giants may commit to fostering an open and inclusive environment, but they

Alt-right: (short for "alternative right") is a segment of the white supremacist movement consisting of a loose network of racists and anti-Semites who reject mainstream conservatism in favor of politics that embrace implicit or explicit racist, anti-Semitic and white supremacist ideology.

Alt-light (also: alt-lite; the New Right): A grouping which holds many of the same views as the alt-right, but seeks to differentiate itself from white supremacist ideology.

("From Alt Right to Alt Lite: Naming the Hate", *Anti-Defamation League* https://www.adl.org/resources/backgrounders/from-alt-right-to-alt-lite-naming-the-hate)

cannot guarantee that their engineers, coders and decision-makers are free from bias. It is therefore highly improbable that multinational technology corporations or any of their smaller counterparts can eliminate bias by simply firing their most outspoken, politically-incorrect dissenters.

2. Broad, ambiguous principles with uneven or non-existent implementation are dangerous.

Ambiguity in the framing of norms, paired with hypocrisy in their practice—especially by influential states and other norm leaders—weakens the core principle that underwrites them. This can lead to the death of the norm.

There appears to be no dearth of declarations of commitment to AI principles of safety, social benefit, and transparency. At least a dozen countries have issued their respective national AI strategies; corporations like Google and Microsoft have publicised their own documents; and many other groups have done similarly, including the Communist Party-backed Beijing Academy of Artificial Intelligence (BAAI), and the Organisation for Economic Cooperation and Development (OECD).[19] However, mere statements of principles will not ensure that they will in fact be implemented.

For instance, the BAAI, established in November 2018 with support from the Ministry of Science and Technology, called for a harmonised set of AI governance frameworks built through international cooperation and pledged to serve all of humanity under the Beijing AI Principles.[20] Through a Western liberal lens, this may seem incompatible with the Chinese government's unethical practices: after all, there are numerous reports of the CCP's panopticon-like programmes, with oppressive regimes of surveillance and suppression targeting the Uyghurs, a Muslim minority group, and widespread use of facial recognition to nudge citizens to behave the way the state wants them to.[21] However, in the eyes of the CCP, this is not a contradiction. Liu Xiaoming, Ambassador of China to the UK has defended the mass surveillance and extralegal detention of Uyghurs, using the rhetoric of "ethnic harmony".[22] It should come as no surprise then that the AI principles are silent on the military and policing applications of AI.

These inconsistencies are seen not only in non-democracy states. Japan's 2017 *Draft AI R&D Guidelines for International Discussions*, for instance, commit to nine principles, including accountability, privacy and ethics.[23] Just one year prior, the Supreme Court dismissed a case against mass surveillance of Muslims, even those on whom the government had no proof of illegal activity. Under the programme, Tokyo Metropolitan Police obtained financial and personal data of over 70,000 Muslim residents, all without consent.[24]

India's *National Strategy for Artificial Intelligence,* for its part, similarly places an emphasis on "Responsible AI", but in the absence of robust data laws and other core protections, several applications of AI remain in a regulatory grey area and susceptible to misuse. For instance, facial recognition trials under *DigiYatra,* a system which will authenticate against Aadhaar by creating a 'token', have already begun at major airports across the country. Yet, concerns about data protection raised under Aadhaar remain

unresolved and there is little clarity on pathways for a true opt-out, should a passenger so choose.[25]

For any set of principles to become norms, norm entrepreneurs must "convince a critical mass of norm leaders" to embrace and then perpetuate these new norms.[26] The espousal of basic AI norms like diversity, transparency, social benefit and safety across different sets of AI principles reflect the beginning of their life cycle. However, for these norms to reach the stage of internalisation, they must meet several conditions, one of which is clarity. The consequences of non-adherence should also be unambiguous.[27] The discrepancy between principle and practice, especially among AI norms leaders like large technology firms and great "AI race" powers, sets a precedent that weakens the core intention of the norm. Nascent AI norms entangled in these dynamics will likely die a premature death.

3. The same applications of AI lend to different concerns when used in varying contexts.

There can be no "catch-all" norm or regulation in the field of AI because the consequences are heavily influenced by the context in which it is used.

AI is, perhaps more than any technology before it, exploitable to a fault; there is a veritable cornucopia of open-source AI resources like TensorFlow and Microsoft's Cognitive Toolkit. AI also finds applicability in a wide gamut of industries and contexts: from financial services to machine maintenance, from personal voice assistants to military decision-support systems. While similar AI applications are used across different fields, not all of them are underpinned by the same concerns.

Facial recognition is a germane case. India's DigiYatra project was announced in June 2018 with the mission to "develop a digital ecosystem that will deliver Indian customers a seamless, consistent and paperless service experience at every touch point of their journey."[28] As part of this project, the Hyderabad international airport launched facial recognition trials for certain domestic routes for the month of July 2019, with 180 volunteers registering for paperless entry on day one.[29] In the same month, the London Metropolitan Police were hit with a report from the University of Essex calling into question the accuracy, legality and ethicality of the Met's live facial recognition technology.[30] Metropolitan Police Federation Chairman Ken Marsh dismissed the report's findings; in the same media interview, he expressed admiration for the Chinese government's surveillance system, which many have described as intrusive.[31]

AI experts and civil society have raised concerns about facial recognition technology's broader implications. These include the risks posed by storing all of one's biometric details on a database which could be vulnerable, as well as this technology's unprecedented infringement of people's privacy.[32]

Yet, the implications are different in the two contexts described above. In Hyderabad, the use-case risk is that of inconvenience: facial recognition is not always accurate, which may lead to the very same delays and queues that DigiYatra purportedly aims to solve. In

London, meanwhile, the same inaccuracies of facial recognition technology can lead to potential violations of civil and human rights: for instance, the arrest and detention of innocent people, or racial profiling.

Points of contention between different applications of the same AI technologies have and will continue to exist. While some actors have sought to avoid these challenges by banning the technology entirely, others are pushing forward with facial recognition applications, drawn by the allure of convenience. In varied use cases such as the ones illustrated in this section, it may be useful to ask the following questions for each case: Does it have the potential to benefit everyone equally or does it disproportionately harm certain groups? Are the risks worth the rewards? If they are, what norms should be propagated to mitigate risks?

4. In defense of bias

Bias need not always have a negative connotation. Some bias is essential in any set of algorithms deployed in a hierarchical society to alleviate existing inequities.

AI is widely expected to transcend the limitations of human cognition, therefore functioning independent of social hierarchies. However, this kind of framing tends to completely disregard the fact that AI will ultimately be deployed in these same stratified societies.

A case in point is Amazon's AI-based recruitment system, trained on years of hiring data, which ended up inadvertently exposing the rampant biases in Amazon's hiring process. The AI-driven recruiting engine was trained on a decade's worth of resumes received by Amazon. AI tends to perform best when the goals are clearly outlined, and the recruiting engine excelled at what it was programmed to do: detecting patterns in Amazon's hiring process to determine their ideal candidate. "In effect, Amazon's system," as a *Reuters* report summarised, "taught itself that male candidates were preferable."[33] Since the majority of resumes were from men, the system downgraded resumes with women-specific terms and language. While Amazon attempted to neutralise this bias, they ultimately had to scrap the project.

The lesson appears to be that removing all bias is not the answer: ignoring the existence of hierarchies will not change the fact that they do exist. In fact, it can be argued that 'unbiased' AI will only perpetuate the status quo. As an illustration, one can reframe Amazon's recruitment algorithm such that it is free from bias and only picks resumes based on experience. It sees two resumes, both with seven years of experience but in the second, there is a one-year employment gap. The algorithm disqualifies the second resume; the second resume belongs to an accomplished woman who had to leave the workforce for a year to care for her newborn. There are several factors that, even if assessed by an algorithm trained to be completely unbiased, would essentially serve as proxies for the social categorisations the AI is supposedly 'blind' to.

The above example is a lengthy way of stressing an oft-ignored point: equality does not ensure equity. There is therefore a fundamental contradiction in the argument that removal of bias will result in equal

outcomes for all. Societies will remain flawed for the near future: the true challenge is in crafting algorithms to move the world closer to its aspirations. The way to do that would be by incorporating some 'bias' that helps alleviate some of the deepest fissures that exist.

CONCLUSION

The task of creating universal AI norms is daunting: it needs to contend with global upheavals that have shaken consumer trust and widened inter-state and inter-group fissures. Furthermore, norms would also need to account not only for the technology itself but also the various contexts they will be used in, all of which would result in radically different ethical and legal concerns.

As nations around the world struggle to adapt to intelligent systems and mitigate their harmful effects, they may need to re-evaluate the scope of the challenge. The basic principles that need to drive responsible development of AI may seem self-evident, but as Isaac Asimov has said, "It is the obvious which is so difficult to see most of the time. People say 'It's as plain as the nose on your face. But how much of the nose on your face can you see, unless someone holds a mirror up to you?" Beyond consequentialist considerations like "AI will take our jobs" v. "AI will liberate us to pursue our full potential", rule- and norms-building around AI is an opportunity to introspect on what human values we want to impart to a technology that will shape society and grow to permeate every aspect of our lives.ORF

ABOUT THE AUTHOR

Trisha Ray is a Junior Fellow with ORF's Cyber Initiative.

ENDNOTES

1. James Vincent, "China and the US are battling to become the world's first AI superpower", *The Verge*, August 3, 2017. https://www.theverge.com/2017/8/3/16007736/china-us-ai-artificial-intelligence
2. Oren Etzioni, "How to Regulate Artificial Intelligence", *New York Times*, September 2, 2017. https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html
3. The Federal Government of Germany, *Artificial Intelligence Strategy* (November 2018) https://www.ki-strategie-deutschland.de/home.html?file=files/downloads/Nationale_KI-Strategie_engl.pdf. European Commission High Level Group on AI, *Ethics guidelines for trustworthy AI,* (April 8, 2019), https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. "OpenAI Charter", OpenAI. https://openai.com/charter/, "Artificial Intelligence at Google: Our Principles", Google AI. https://ai.google/principles/. "OECD Principles on AI", OECD. https://www.oecd.org/going-digital/ai/principles/
4. Ivana Kottasova, "Europe is making AI rules now to avoid a new tech crisis", *CNN Business*. April 8, 2019. https://edition.cnn.com/2019/04/08/tech/ai-guidelines-eu/index.html
5. US Director of National Intelligence, *Assessing Russian Activities and Intentions in Recent US Elections*, ICA 2017-01D, January 6, 2017. https://www.dni.gov/files/documents/ICA_2017_01.pdf . Nicole Kobie, "The complicated truth about China's social credit system", *Wired*, June 7, 2019. https://www.wired.co.uk/article/china-social-credit-system-explained
6. Kent Walker, "An external advisory council to help advance the responsible development of AI", *Google*, March 26, 2019. https://www.blog.google/technology/ai/external-advisory-council-help-advance-responsible-development-ai/
7. Alessandro Acquisti, Twitter post, March 31, 2:37 a.m. https://twitter.com/ssnstudy/status/1112099054551515138. Julia Angwin, "Google Has Quietly Dropped Ban on Personally Identifiable Web Tracking", *ProPublica*, October 21, 2016. US Congress, Senate, Committee on the Judiciary, Subcommittee on the Constitution, *Google and Censorship through Search Engines: Hearing before the Senate Judiciary Subcommittee on the Constitution*, Testimony of Karan Bhatia, July 16, 2019. https://www.judiciary.senate.gov/imo/media/doc/Bhatia%20Testimony.pdf
8. Googlers Against Transphobia, "Googlers Against Transphobia and Hate", *The Medium,* April 1, 2019. https://medium.com/@against.transphobia/googlers-against-transphobia-and-hate-b1b0a5dbf76
9. Nick Statt, "Google dissolves AI ethics board just one week after forming it", *The Verge*. April 4, 2019. https://www.theverge.com/2019/4/4/18296113/google-ai-ethics-board-ends-controversy-kay-coles-james-heritage-foundation
10. "Morning Consult National Tracking Poll #170401", *Morning Consult/Politico,* April 1, 2017. https://morningconsult.com/wp-content/uploads/2017/04/170401_crosstabs_Politico_v3_AG.pdf
11. Smita Sinha, "Annual Consumer Survey on Data Privacy in India", *Analytics India*, May 25, 2018. https://www.analyticsindiamag.com/annual-consumer-survey-on-data-privacy-in-india-2018/
12. Kenneth Corbin, "Warren Wants to Break Up Amazon, Facebook, Google", *Forbes*, March 8, 2019. https://www.forbes.com/sites/kennethcorbin/2019/03/08/warren-wants-to-break-up-amazon-facebook-google/. "Antitrust: Commission fines Google €1.49 billion for abusive practices in online advertising", *European Commission*, March 20, 2019.

http://europa.eu/rapid/press-release_IP-19-1770_en.htm. Ministry of Electronics and Information Technology, *The Personal Data Protection Bill, 2018* (2018). https://www.meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf

13. James Damore, "Google's Ideological Echo Chamber", July 2017. https://www.documentcloud.org/documents/3914586-Googles-Ideological-Echo-Chamber.html
14. Lela Moore and Briana Milord, "'Thought Bullies' or Right Move: A Divide Over James Damore Firing", *The New York Times*, August 10, 2017. https://www.nytimes.com/2017/08/10/technology/google-james-damore-memo-commentary.html. Abby Ohlheiser, "How James Damore went from Google employee to right-wing Internet hero", *The Washington Post,* August 12, 2017. https://www.washingtonpost.com/news/the-intersect/wp/2017/08/12/how-james-damore-went-from-google-employee-to-right-wing-internet-hero/
15. Jane McGonical, "Our Puny Human Brains Are Terrible at Thinking About the Future", *Slate,* April 13, 2017. https://slate.com/technology/2017/04/why-people-are-so-bad-at-thinking-about-the-future.html
16. "Bounded Ethicality", Ethics Unwrapped, University of Texas. https://ethicsunwrapped.utexas.edu/glossary/bounded-ethicality
17. Sheera Frenkel, Nicholas Confessore, Cecilia Kang et al, "Delay, Deny and Deflect: How Facebook's Leaders Fought Through Crisis", *The New York Times*, November 14, 2018 https://www.nytimes.com/2018/11/14/technology/facebook-data-russia-election-racism.html
18. Sarah Kessler, "The History of Advertising on Facebook", *Mashable,* June 28, 2011. https://mashable.com/2011/06/28/facebook-advertising-infographic/
19. Samir Saran, Nikhila Natarajan and Madhulika Srikumar, *In Pursuit of Autonomy: AI and National Strategies* (November 16, 2018). https://www.orfonline.org/research/in-pursuit-of-autonomy-ai-and-national-strategies/. "Artificial Intelligence at Google: Our Principles", Google AI. https://ai.google/principles/. "Microsoft AI Principles", Microsoft. https://www.microsoft.com/en-us/ai/our-approach-to-ai. Beijing Academy of Artificial Intelligence, *Beijing AI Principles* (May 28, 2019). https://www.baai.ac.cn/blog/beijing-ai-principles. "OECD Principles on AI", OECD. https://www.oecd.org/going-digital/ai/principles/
20. Beijing Academy of Artificial Intelligence, *Beijing AI Principles* (May 28, 2019) https://www.baai.ac.cn/blog/beijing-ai-principles
21. "China's Algorithms of Repression", *Human Rights Watch,* May 1, 2019. https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance. "How China Is Using "Social Credit Scores" to Reward and Punish Its Citizens", *Time*, January 16, 2019 https://time.com/collection/davos-2019/5502592/china-social-credit-score/
22. Liu Xiaoming, "Harmony in Xinjiang is based on three principles", *Financial Times*, August 21, 2018. https://www.ft.com/content/05a81682-a219-11e8-85da-eeb7a9ce36e4
23. "Draft AI R&D Guidelines for International Discussions", *The Conference toward AI Network Society,* July 28, 2017. http://www.soumu.go.jp/main_content/000507517.pdf
24. "Spying on Muslims in Tokyo and New York — "Necessary and Unavoidable"?", *The Asia-Pacific Journal*, Volume 14, Issue 18, Number 2, September 15, 2016. https://apjjf.org/—Asia-Pacific-Journal-Report/4956/article.pdf

25. Vidushi Marda, Srinivas Kodali, "The Privacy Cost of Digi Yatra's Travel Promise", *Huffpost*, July 16, 2019. https://www.huffingtonpost.in/entry/digi-yatra-face-recognition-hyderabad-airport_in_5d2d9725e4b085eda5a15c28
26. Martha Finnemore and Kathryn Sikkink, "International Norm Dynamics and Political Change", *International Organization* 52(4) (1998), 887–917.
27. Ibid
28. "'DigiYatra'- A New Digital Experience for Air Travellers", *Press Information Bureau, Government of India*, June 8, 2017. http://pib.nic.in/newsite/PrintRelease.aspx?relid=165513
29. "Hyderabad airport launches Face Recognition system for entry", *The Hindu*, July 2, 2019. https://www.thehindubusinessline.com/economy/logistics/hyderabad-airport-launches-face-recognition-system-for-entry/article28261687.ece
30. Pete Fussey and Daragh Murray, "Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology", *The Human Right, Big Data and Technology Project, University of Essex,* July 2019. https://48ba3m4eh2bf2sksp43rq8kk-wpengine.netdna-ssl.com/wp-content/uploads/2019/07/London-Met-Police-Trial-of-Facial-Recognition-Tech-Report.pdf
31. Ken Marsh, "Should facial recognition software be used?", Interview by Sonia Watson and Ben Fryer, *Breakfast with Ben and Sonia, BBC*, July 5, 2019. https://www.bbc.co.uk/sounds/play/p07d51mr
32. "China's Algorithms of Repression", *Human Rights Watch,* May 1, 2019. https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance. "How China Is Using "Social Credit Scores" to Reward and Punish Its Citizens", *Time*, January 16, 2019 https://time.com/collection/davos-2019/5502592/china-social-credit-score/
33. Ron Hurtibise, "With Facial Recognition Comes Convenience, But What About Privacy?", *Government Technology*, April 17, 2019. https://www.govtech.com/security/With-Facial-Recognition-Comes-Convenience-But-What-About-Privacy.html. Vidushi Marda, Srinivas Kodali, "The Privacy Cost of Digi Yatra's Travel Promise", *Huffpost*, July 16, 2019. https://www.huffingtonpost.in/entry/digi-yatra-face-recognition-hyderabad-airport_in_5d2d9725e4b085eda5a15c28
34. Jeffrey Dastin. "Amazon scraps secret AI recruiting tool that showed bias against women". *Reuters*. October 18, 2018. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

ORF OBSERVER RESEARCH FOUNDATION
Ideas • Forums • Leadership • Impact

20, Rouse Avenue Institutional Area, New Delhi - 110 002, INDIA
Ph. : +91-11-35332000 Fax : +91-11-35332005
E-mail: contactus@orfonline.org
Website: www.orfonline.org