The threats of technology to minority and indigenous rights

Michael Caster

The global 'digital divide' continues to prevent ethnic, religious and linguistic minorities and indigenous peoples from accessing the internet and associated information and communication technologies (ICTs) that may support peace, democracy and the promotion of human rights.

Imams and government officials pass under security cameras as they leave the Id Kah Mosque during a government organised trip in Kashgar, Xinjiang Uighur Autonomous Region, China. *REUTERS / Ben Blanchard*

Sadly, patterns of exclusion and discrimination in everyday life are mirrored online; the United Nations (UN) reports that nearly half the world's population is not connected to the internet,[1] while the Organisation for Economic Co-operation and Development (OECD) estimates that the proportion of women using the internet is 12 per cent lower than that of men.[2]

Globally, marginalized ethnic groups have worse internet access than dominant ethnicities in the same country.[3] This remains the case despite the UN Human Rights Council (HRC) having stated back in 2011: 'Given that the Internet has become an indispensable tool for realizing a range of human rights, combating inequality, and accelerating development and human progress, ensuring universal access to the Internet should be a priority for all States.'[4]

While the internet and ICTs have great potential to challenge entrenched discrimination, the limited access of minorities and indigenous peoples to these technologies threatens to exacerbate their situation further. This is why abusive governments, especially across Asia, have increasingly turned to internet shutdowns to target certain ethnic and religious communities, taking away their freedom of expression and ability to document and disseminate evidence of ongoing human rights abuses. Intentionally shutting down or restricting access to the internet can in and of itself be a human rights violation, while also causing the proliferation of other rights abuses as it prevents victims from documenting and sharing them online, from Cameroon to West Papua. This can also complicate future attempts at accountability. In 2019 alone, the digital rights organization Access Now documented some 213 internet shutdowns. This includes a 47 per cent increase across Africa, with Ethiopia identified as one of the worst offenders. However, India alone accounted for more than half of the total in 2019, with a single shutdown in Indian-controlled Kashmir lasting for nearly six months.

Even where access to the internet and other ICTs is not arbitrarily denied, minorities and indigenous peoples are frequently targets of online hate speech and sophisticated surveillance technologies. As noted in 2019 by the UN Special Rapporteur on freedom of opinion and expression,

'The prevalence of online hate poses challenges to everyone, first and foremost the marginalized individuals who are its principal targets.'[5] Examples include parts of Europe where online content vilifying refugees and migrants has been correlated to physical attacks against them, or the spread of anti-Rohingya speech on Facebook in Myanmar which has been tied to acts of genocide.'

[1] UN ITU, 'Statistics', https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx

[2] OECD, *Bridging the Digital Gender Divide*, 2018, p. 25.

[3] Weidmann, N.B., Benitez-Baleato, S., Hunziker, P., Glatz, E. and Dimitropoulos, X., 'Digital discrimination: political bias in internet service provision across ethnic groups', *Science*, 353(6304), pp. 1151–5.

[4] UN HRC, Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Frank La Rue, A/HRC/17/27, 16 May 2011.

[5] OHCHR, 'Governments and internet companies fail to meet challenges of online hate – UN expert', 21 October 2019.

Many such concerns remain unresolved due to the lack of universally accepted obligations by public and private actors, issues of transparency and the poor implementation of existing human rights law in the digital age.

The impact of new technologies can also be more insidious. For example, not only is big data and machine learning allowing for the automation of human decision-making in governance and criminal justice – a situation that risks replicating historical injustices through algorithmic bias – but it is also increasingly leading to labour displacement, impacting particularly on minority communities. In the United States (US), a 2019 study by the Brookings think-tank noted that the 'average current-task automation potential' among Hispanics, Native Americans and black Americans was 47 per cent, 45 per cent and 44 per cent respectively, compared to 40 per cent among white Americans.[6]

Social media and the internet

> *'The last decade has seen minorities around the world facing new and growing threats, fuelled by hate and bigotry being spewed through social media platforms… This has contributed to the rise of violent extremist groups and to a dramatic increase in many countries of hate crimes targeting religious, ethnic and other minorities, including migrants.'*

UN Special Rapporteur on minority issues Fernand de Varennes, 2020[7]

The role of social media in spreading hate speech is compounded when social media is effectively your only access to the internet, such as with Free Basics, a Facebook product providing free limited internet access in developing markets. Myanmar is emblematic, as noted in the UN independent international fact-finding mission: 'Facebook has been a useful instrument for those seeking to spread hate, in a context where, for most users, Facebook is the Internet.'[8] Online, hate speech against Rohingya is rife, including comparisons of Rohingya to animals, accusations that Rohingya stage human rights abuses against themselves, and direct threats against them. According to one study, 1 in 10 of the social media posts by politicians of the Arakan National Party (ANP) contained hate speech. The ANP is the main party representing the dominant Rakhine ethnic group in Rakhine State, where most Rohingya lived prior to their mass displacement in 2017-18. The most popular hate messages by members of the Rakhine State parliament received 3,400 reactions or were shared up to 9,500 times.[9]

[6] Muro, M., Maxim, R. and Whiton, J., *Automation and Artificial Intelligence: How Machines Are Affecting People and Places,* Washington, DC, Brookings, 2019, p. 44.

[7] OHCHR, 'UN expert denounces the propagation of hate speech through social media', 27 February 2020.

[8] OHCHR, Report of the Independent International Fact-finding Mission on Myanmar, 2018, p. 14.

[9] Rajagopalan, M., Vo, L.T. and Soe, A.N., 'How Facebook failed the Rohingya in Myanmar', *Buzzfeed*, 27 August 2018.

According to one study, popular hate messages by members of the Rakhine State parliament in Myanmar received **3,400 reactions or were shared up to 9,500 times on social media.**

In response to such trends, in late 2018, Facebook admitted they had not done enough to prevent their ‘platform from being used to foment division and incite offline violence’, and vowed to do more to counter hate speech.

Meanwhile, in the name of combating ‘fake news’ or protecting national security, the Myanmar government has also at times blocked Facebook or shut down internet access in entire minority townships in Rakhine and Chin states. Under international law, freedom of expression and information may only be restricted under narrowly defined circumstances: namely, restrictions must be prescribed by law with sufficient precision to enable regulation; they must pursue a legitimate aim respecting other rights such as non-discrimination; and be necessary and proportionate. Responding to the shutdown, the non-governmental organization (NGO) Article 19 found that Myanmar failed to meet these basic requirements. Such measures arbitrarily restrict freedom of expression, while also making it harder to document and disseminate evidence of human rights abuses against Rohingya and other minority populations. This can have long-term impacts on accountability. Here, arguably, the silencing of human rights abuses was not an unintended consequence but the specific aim of the shutdown. Similar problems occurred in Sri Lanka following the blocking of social media, purportedly to prevent the spread of rumours, after the 2019 Easter massacre.

In India, there have been accusations that social media have been weaponized against non-Hindu minorities, leading to communal violence. This is especially the case with WhatsApp, which has over 400 million monthly active users in India. On WhatsApp, Facebook and other platforms, there has been a reported increase in the spread of hate speech and disinformation portraying Muslim citizens as terrorists or rapists, or accusing them of plotting genocide against the Hindu majority. Such is the Hindu nationalist sentiment influencing the recent Citizenship Amendment Act, discriminatory legislation that favours migrants from certain religious communities (Hindu, Sikh, Buddhist, Jain, Parsi and Christian) for fast-tracked citizenship while conspicuously excluding Muslims. Its passage in December 2019 was accompanied by protests and communal violence. Videos of predominantly Muslim minorities being beaten have been shared via WhatsApp, an activity that has been compared to lynchings. In response, WhatsApp has limited the number of times a message can be forwarded, first to 20 and now to 5, but with WhatsApp group sizes of up to 256 people such content, even forwarded only 5 times, could still reach nearly 1,300 people.[10] This is a reminder that technology does not exist in a vacuum, and merely curbing technology without addressing the underlying contexts of oppression is unlikely to have a significant impact. Such curbs may sound appealing, but they raise fundamental digital rights concerns. WhatsApp communications are end-to-end encrypted and the implementation of such a law would require removal of this protection, setting a dangerous precedent.

Encrypted and anonymous communication is important to protect the right to privacy and freedom of expression online, but it is also crucial for protecting vulnerable populations, such as ethnic, religious or sexual minorities, against arbitrary and unlawful interference or attacks. While India is not the only government to challenge encryption through legislation, private companies are also rolling out malware capable of attacking user privacy. In 2019, WhatsApp filed a lawsuit against Israeli spyware company NSO Group over a hack of 1,400 users, from Indian journalists to Rwandan human rights defenders. It has also been pointed out that Facebook's acquisition of WhatsApp, and its plans to integrate Instagram and WhatsApp with its own messaging service, have given rise to new digital security concerns in addition to the potential insecurities created by publicly shared hateful content on such platforms.

While stricter content moderation standards might seem to be an obvious solution, this also poses challenges, especially when companies are not transparent about what is removed or how this is done. One effort to broadly improve social media in this regard, the 2018 Santa Clara Principles on Transparency and Accountability in Content Moderation, were put forward, calling on companies to publish the number of posts removed and accounts suspended, notify users of the reasons why their content is removed or accounts suspended, and to ensure effective means of appeal. But addressing hate speech online is not as simple as just removing hateful content or flagging abusive accounts. Grasping cultural, religious or linguistic nuances requires linguistic fluency, but the promise of fluency in local languages can also come with local anti-minority biases. Different platforms and jurisdictions have their own policies and inconsistencies. In the US, for example, Facebook's efforts to remove hate speech have also inadvertently censored minority groups using the platform to call out racism or create dialogue. In some countries, laws intended to protect minorities from online hate speech have instead engendered censorship and risked violating other rights.

Germany, in response to the role of hate speech in the early normalization of Nazi atrocities against Jews, Roma and other minorities during the Second World War, has some of the harshest hate speech laws. Since 2018, the Network Enforcement Act (NetzDG) requires social media companies like Facebook, Twitter and YouTube to remove 'illegal content' within 24 hours or risk fines of up to €50 million. In 2019, Australia also passed a law to penalize social

[10] Kastrenakes, J., 'WhatsApp limits message forwarding in fight against misinformation', *The Verge*, 21 January 2019.

Since 2018, the Network Enforcement Act (NetzDG) requires social media companies to remove ‘illegal content’ **within 24 hours or risk fines of up to €50 million**.

media platforms for not removing certain content, carrying the potential risk of up to three years in prison for executives of companies who fail to do so. The United Kingdom (UK) is discussing similar legislation to combat hate speech, disinformation, cyberstalking and terrorist activity by creating a single regulator that can also penalize social media platforms. Compare these laws to Section 320 of the US Communication Decency Act, which holds that no social media platform shall be held liable for content provided by someone else, an important protection for free speech in that social media companies that may be held liable for speech on their platforms are likely to over-censor. While Germany and Australia have a generally functioning rule of law, some governments without independent judiciaries are also turning to laws like NetzDG to inspire their own regulations, including Russia, Belarus, Singapore, Vietnam and the Philippines.

In February 2020, Ethiopia, recently transitioning out of authoritarian rule and a past of wielding the law to detain and silence dissent, passed a law against hate speech that will punish online dissemination of hate speech or disinformation with up to three years in prison. With over 90 distinct ethnic groups, Ethiopia has a history of marginalizing minority communities such as Oromo and Amhara, and in March 2020 the government shut down the internet in much of the Oromia region, amid reports of human rights abuses against the armed Oromo Liberation Front. Displacement of Ethiopian indigenous peoples, largely in the Gambella and Lower Omo regions, is also common. Hate speech has admittedly fuelled inter-ethnic violence, but without robust oversight and due process, instead of protecting such marginalized communities, the new anti-hate speech law may have a chilling effect on freedom of expression and inter-ethnic dialogue.

Nigeria is also considering harsh legislation that would allow authorities to shut down the internet, limit social media access, and make criticism of the government punishable with up to three years’ imprisonment, and even, in some cases, impose life imprisonment or the death penalty for hate speech. Nigeria has a diverse population of some 250 distinct ethno-linguistic groups. It is a country that has long witnessed numerous conflicts over varying political and economic interests. For instance, tensions over land and water between settled farmers and nomadic herders in the Middle Belt have led to over 10,000 people being killed in the last decade alone. Ogoni and other minorities in the southern Niger Delta region have particularly faced persecution in connection with oil and gas extraction. If new laws intended to crack down on inter-ethnic violence or hate speech are not properly monitored, they may silence documentation and

SURVEILLANCE AND DIGITAL FREEDOMS

Ethiopia: recently passed a law against hate speech that will punish online dissemination of hate speech or disinformation with up to three years in prison.

Nigeria: is considering harsh legislation that would allow authorities to shut down the internet, limit social media access, and make criticism of the government punishable with up to three years' imprisonment.

Brazil: in forming its council to counter fake news included the military and domestic intelligence services, both of which have a record of harassing, silencing and crushing minority and indigenous communities.

dissemination of such rights abuses without addressing the root causes of intolerance and discrimination.

Elsewhere, laws developed in the name of combating 'fake news' and online disinformation have been proposed or enacted in multiple countries, with alarming ramifications for human rights. When Brazil formed its council to counter fake news, it included the military and domestic intelligence services, both of which have a record of harassing, silencing and crushing minority and indigenous communities. Under such legislation, for example, indigenous rights defenders documenting land-grabbing could be criminalized if their campaigns become labelled as fake news. Recognizing such global concerns, the Organization of American States (OAS), the African Commission on Human and Peoples' Rights, the Organization for Security and Co-operation in Europe and others put forward in 2017 the 'Joint Declaration on Freedom of Expression

Victor holds up a leaf coated in oil as he stands in an oil polluted fishpond in Ogoniland, Niger Delta. *George Osodi* and “Fake News,” Disinformation and Propaganda’ as a guideline for a rights-based approach to managing potentially harmful content.

Moving beyond individual social media platforms, there are growing concerns around what has been termed cyberbalkanization or internet balkanization. This notion relates to some of the localized cyber laws noted above but goes well beyond in its theorization of internet ecosystems. In September 2018, former Google chief executive Eric Schmidt put forward the idea, during a meeting with a venture capital firm, that in the next 10 to 15 years the internet would be split between China and the US. China, after all, has perfected centralized internet control under the Great Firewall and an ever increasing armada of artificial intelligence (AI)-supported censorship applications so that banned topics, such as discussion of the persecution of Uyghurs and Tibetans, is not only criminalized but wiped from the Chinese internet and social media platforms. In China, the internet is not a reflection of reality but of the propaganda of the ruling Communist Party, and all the characterization or masking of minority persecution that comes with it. China calls it a ‘sovereign internet’, but such ideas mean the proliferation of human rights abuses online. It is little wonder that other authoritarian states are following suit, and in 2019 Russia adopted its own ‘Sovereign Internet Law’ based on the China model. Meanwhile, as Iranian-Canadian media scholar Hossein Derakhshan points out, the European Union (EU)’s General Data Protection Regulation (GDPR) and related laws on hate speech, privacy and copyright are essentially turning the EU-based internet into its own separate legal sphere. Signs point to a three-tiered internet in the future – the US, China and Europe – with potentially vastly different risks and protection regimes for minority rights in the digital age.

Surveillance and digital freedoms

‘Surveillance tools can interfere with human rights, from the right to privacy and freedom of expression to rights of association and assembly, religious belief, non-discrimination, and public participation.’

UN Special Rapporteur on the freedom of expression David Kaye, 2019[11]

In 2013, the UN General Assembly adopted a resolution on the right to privacy in the digital age, which expressed deep concerns over the negative impact that surveillance and the mass collection of personal data can have on human rights. Nowhere is this more pronounced than in China.

China has perfected sophisticated surveillance systems designed to profile ethnic and religious minorities, namely Uyghur, Kazakh, Kyrgyz and Hui Muslims in the Xinjiang Uyghur Autonomous Region. This surveillance is part of the mass extra-judicial internment, disappearance, torture and forced labour in Xinjiang that has led to widespread calls for a UN-led investigation. One of the main systems by which China enforces

[11] OHCHR, ‘UN expert calls for immediate moratorium on the sale, transfer and use of surveillance tools’, 25 June 2019.

the widespread surveillance of some 13 million regional Turkic Muslims is the Integrated Joint Operations Platform (IJOP), which Human Rights Watch (HRW) revealed in 2019 is used by authorities to collect and centralize massive amounts of personal information, from hair colour and height to private religious and cultural beliefs and whether family members have studied abroad.

Such platforms utilize AI to identify people through facial or voice recognition, and other machine-learning algorithms based on the mass forced collection of biometric data, such as DNA, fingerprints, iris scans and blood samples. When combined with ubiquitous checkpoints, IJOP also functions as a virtual fence, restricting freedom of movement in the real world. No longer confined to Xinjiang, the police in China have expanded on these technologies to target Uyghurs living across the country. According to a report by the *New York Times*, in April 2019 alone police in one central Chinese city ran facial recognition surveillance to determine if residents were Uyghurs some 500,000 times. Such technology is on the rise in China.[12]

But China is also a world-leading source for AI surveillance to other countries. As recently reported by the Carnegie Endowment for International Peace, Chinese technology firms such as Huawei, Hikvision, Dahua and ZTE supply AI surveillance technologies to some 63 countries, 36 of which are members of China's Belt and Road Initiative. The fact that these products are often marketed with the help of loans from the Chinese government, including to countries which might otherwise not have the resources to purchase them, raises 'troubling questions about the extent to which the Chinese government is subsidizing the purchase of advanced repressive technology'.[13] In light of the human rights violations perpetrated with Chinese surveillance technology, it is furthermore concerning that companies such as ZTE, Dahua and others are communicating with the UN International Telecommunication Union (ITU) to shape new international standards on facial recognition surveillance.

Governments with abysmal human rights records are not the only ones employing or abusing surveillance technologies against ethnic and religious minority citizens. In the UK, following the 2011 London riots, the Metropolitan Police launched the Gangs Matrix program. A system utilizing AI and machine learning to compile a database of gang members, it has been criticized by Amnesty UK as 'a racially discriminatory system that stigmatises young black men for the music they listen to or their behaviour on social media'. According to a 2019 Freedom of Information Request obtained by *WIRED*, some 80 per cent are listed as 'African-Caribbean', with a further 12 per cent from other ethnic minority groups, while only the remaining 8 per cent are listed as 'white European'. Some are

[12] Mozur, P., 'One month, 500,000 face scans: how China is using AI to profile a minority', *New York Times*, 14 April 2019.

[13] Feldstein, S., *The Global Expansion of AI Surveillance*, Washington, DC, Carnegie Endowment for International Peace, September 2019.

as young as 12.[14] Since its inception, the database has listed around 7,000 people, and once someone is on the Matrix, finding out why or getting their name removed can be extremely difficult. But, in a victory for privacy and anti-discrimination advocates, several hundred names were removed from the Matrix in early 2020, correcting for ethnic bias and violations of data protection laws. Similarly, for many years following the 11 September 2001 attacks, the New York City Police Department (NYPD) engaged in a Muslim Surveillance Program that combined digital surveillance with informants and other types of physical surveillance, giving rise to numerous human rights concerns over the discriminatory targeting and stigmatization of religious minorities.

In Canada, police networks, the Canadian Security Intelligence Service (CSIS) and other government agencies have subjected indigenous rights defenders to abusive surveillance and hacking, often by labelling them 'multi-issue extremists'. This charge is largely in response to indigenous protests against oil and gas pipelines, hydroelectric dams, mining operations and other extractive industries due to environmental concerns and encroachment on indigenous land. In some cases, CSIS has worked directly with energy companies to conduct surveillance of indigenous peoples. In others, police surveillance has been clearly excessive, such as a 16-month undercover operation in Saskatchewan Province to catch an indigenous man accused of illegally selling 90 Canadian dollars' worth of fish.

These examples demonstrate that both open and repressive governments are engaged in surveillance practices that raise human rights concerns. As such, in responding to *Privacy International v. the United Kingdom*, a current case of government surveillance before the European Court of Human Rights, Article 19 and the Electronic Frontiers Foundation (EFF) among others point out that government surveillance, including hacking, has a 'chilling effect' on online expression, contributing to self-censorship or preventing them from organizing or supporting protests. It has also been shown to particularly impact vulnerable groups, members of which may be fearful of reporting online abuse.

Border crossings have also become hotspots for automated surveillance. The EU has piloted an AI-driven facial recognition lie-detector video surveillance border control system in Hungary, Greece and Latvia called iBorderCtrl. Based on the contested theory of 'affect recognition science', iBorderCtrl replaces human border guards with a video system that scans for facial anomalies while targets answer a series of questions. But the use of this technology at international borders, especially common crossing points for asylum seekers or migrant populations, raises concerns over the potential for bias in facial recognition systems, especially with regard to the analysis of women of colour, cultural-communicative differences, or the inability to distinguish the lingering impact of trauma.

The US has also experimented with 'smart border' technologies along

[14] Yeung, P., 'The grim reality of life under Gangs Matrix, London's controversial predictive policing tool', *WIRED*, 2 April 2019.

the US–Mexico border, relying on automated drones and other surveillance technologies. Such surveillance systems infringe the civil liberties of travellers, immigrants and people living along the border. They also pose other risks: in a 2019 study, researchers in Arizona used geospatial and statistical modelling to show that smart border technologies, instead of preventing undocumented border crossing, merely shifted migration routes to potentially more hazardous terrain, raising the number of migrant deaths in the process. Leading rights groups including EFF and the American Civil Liberties Union (ACLU) have opposed such measures on the grounds they would exacerbate racial and ethnic inequality in policing and immigration enforcement, as well as curbing freedom of expression and the right to privacy.

In 2019, the UN Special Rapporteur on freedom of expression, David Kaye, presented a report on surveillance and human rights before the HRC. He recommended that states impose an immediate moratorium on surveillance tools until proper human rights safeguards are in place and called for an expansion of the Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies to include spyware used to undermine human rights. For private companies, the report recommends that companies should publicly affirm their responsibilities under the UN Guiding Principles on Business and Human Rights to respect 'freedom of expression, privacy and related human rights, and integrate human rights due diligence processes from the earliest stages of product development and throughout their operations'.[15]

In their responses to the tragedy of Covid-19 throughout early 2020, many governments have seized on digital surveillance technologies as part of

[15] HRC, Surveillance and Human Rights: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, A/HRC/41/35, 28 May 2019.

Migrants try to enter Hungary through the border crossing in Horgos, Serbia. The EU has piloted surveillance technology involving a video system that would replace human border guards.

their efforts to contain its spread. While technology can and should play a role in resolving global challenges like the virus, without effective protections and the right to remedy its use also risks serious rights violations – especially for communities which are already discriminated against and marginalized. One such tool has been contact tracing and other mobile app-based tools designed to monitor infected or potentially infected populations, as in many contexts these applications have been developed without taking user privacy or other concerns into account. India is one such example: as Indian activist and writer Arundhati Roy has quipped, 'The coronavirus is a gift to authoritarian states including India.' Indeed, across South Asia governments have been accessing personal data on mobile devices without consent. All of these measures can have wide and long-lasting impacts on the right to privacy, impacting in turn on freedom of movement, association and religion, especially for minorities. Responding to such concerns, in early May 2020 Haroon Baloch of Bytes for All in Pakistan petitioned the Islamabad High Court to disallow such measures. At the time of writing, the case is still pending.

In addition to concerns around surveillance targeting minority communities, there have also been reports of Chinese, other Asian, Roma, Hispanic and other minorities across the world facing hate speech online and physical intimidation due to these groups being accorded blame for the spread of the virus. To make matters worse, minorities and indigenous peoples in many countries may already lack access to medical care due to structural discrimination.

AI and discriminatory bias

People can be biased, but machines are objective – or so many people seem to believe. As machine-learning capabilities improve with more elegant algorithms and big data, the conventional thinking is that the biases or inefficiencies of human-led processes will vanish. But machines are trained by humans and this means that, just as children may learn the ethnic, religious or gender-based biases of their parents or communities, so too can machines develop biases based on their algorithms and datasets. Existing inequalities can be recreated in data, and big data can magnify such inequalities. This is known as algorithmic bias. Organizations like the US-based Algorithmic Justice League have set out to raise awareness of these issues and to mitigate its harms and biases. Confronting this bias is complicated when the algorithms are held in secret by private firms. Another challenge is that even when an algorithm has been corrected for bias against one group, this does not necessarily mean it has corrected for others, especially when discrimination and bias is intersectional. In many cases, from education and employment to policing and criminal sentencing, big data is increasingly influencing our experience in the world. This raises myriad concerns around algorithmic bias.

In 2014 Amazon began to design an AI system to automate parts of the job recruitment process. The algorithm was trained on a dataset based on all the resumés submitted over the previous decade, which also happened to overwhelmingly come from white men.

AI AND DISCRIMINATORY BIAS

China: According to a New York Times report, in April 2019 police in one central Chinese city ran facial recognition surveillance to determine if residents were Uyghurs some 500,000 times.

UK: Following the 2011 London riots, the Metropolitan Police launched the Gangs Matrix program, a system utilizing AI and machine learning to compile a database of gang members. According to a 2019 Freedom of Information Request, some 80 per cent are listed as 'African-Caribbean'.

Canada: Police networks, the Canadian Security Intelligence Service (CSIS) and other government agencies have subjected indigenous rights defenders to abusive surveillance and hacking, often by labelling them 'multi-issue extremists'.

Amazon's hiring machine taught itself to favour this 'baseline' applicant. It is easy to see how, depending on the data inputted, existing inequalities can be replicated in supposedly objective machine learning. For example, if the data draws from majority affluent white male applications, it may score the words 'lacrosse' or 'crew' higher and penalize resumés with words such as 'women's,' as in 'women's chess club captain'. It may equally undervalue extra-curricular activities perhaps more often mentioned among applicants from less affluent and/or minority backgrounds. Although Amazon abandoned its project in 2018, there are a number of automated resumé screening platforms in use on the market today, and certainly not all of them have checked their algorithms for bias. A 2018 survey by LinkedIn revealed that 67 per cent of recruiters and hiring managers globally rely on such tools to 'save time'.

Another example comes from job advertising, as prospective employers turn to the algorithm-based targeting of 'ideal' candidates. Again, depending on the data upon which these machine-driven processes are trained, they can recreate bias. For example, a 2019 study conducted by the technology non-profit Upturn with Northeastern University in Boston and others found that targeted ads on Facebook for grocery cashier positions were shown to audiences of 85 per cent women, while taxi driver jobs were shown to audiences that were 75 per cent black. In a similar case, in 2019 the US Department of Housing and Urban Development (HUD) charged Facebook for violating the Fair Housing Act after it came to light that Facebook user data was being used to influence targeted housing-related advertising that was unlawfully discriminating 'based on race, colour, national origin, religion, familial status, sex and disability'. Training machine learning based on historical employment prejudices or economic and racial housing discrimination ensures their perpetuation. In other words, although such technologies were dreamed up to be disruptive or progressive, relying on supposedly unbiased algorithms to see past

A 2019 study found that targeted ads on Facebook for grocery cashier positions were shown to audiences of **85 per cent women**, while taxi driver jobs were shown to audiences that were **75 per cent black**.

discrimination in the recruitment process, they are just as likely to maintain or reaffirm an unequal status quo.

As seen with London's Gangs Matrix, predictive policing measures cannot be objective when the data they learn from is based on ethnic or other structural and historical biases. As Andrea Nill Sanchez, executive director of the New York University-affiliated AI Now Institute, testified before the European Parliament in February 2020, 'left unchecked, the proliferation of predictive policing risks replicating and amplifying patterns of corrupt, illegal and unethical conduct linked to legacies of discrimination that plague law enforcement agencies across the globe'.

One American company, PredPol, is deployed across the country and offers location-specific predictive policing solutions. Trained from years of recent crime data, it is based on the idea that criminal activity at a certain place is more likely to occur there again and concentrates police activity accordingly. However, this immediately becomes problematic in light of historic over-policing in minority communities. In this case, machine learning based on data from over-policed neighbourhoods feeds an algorithm that predicts the need for more police presence, creating a discriminatory feedback loop. A recent study of predictive policing across England and Wales by the Royal United Services Institute (RUSI) likewise uncovered this problem of replication and amplification of discrimination. In many cases, such as with PredPol and the police departments it partners with, the lack of meaningful transparency between private and public entities makes it increasingly difficult to audit algorithms for bias.

Big data for predictive policing logically gives way to big data for predicting incarceration, with the same concerns of algorithmic bias based on a criminal justice system rife with institutional racism. In the US, pre-trial risk assessments performed by AI are taking place in nearly every state to determine matters such as the likelihood that the accused person will re-offend (known as their 'recidivism risk') or whether they will appear at trial. Such AI-driven decisions can, among other things, impact the chances or terms of bail, sentencing and parole. One such tool, COMPAS by Northpointe, was profiled in a 2016 investigation by ProPublica that showed that while the algorithm was correct over 60 per cent of the time, it also exhibited racial bias when it was wrong. Non-re-offending black defendants were twice as likely to be assigned higher recidivism rates than white defendants, whereas roughly 50 per cent of re-offending white defendants were assigned a lower number. In other words, when it was wrong the algorithm thought black people were more likely, and white people less likely, to commit another crime. This has serious real-world implications on who is imprisoned and for how long, perpetuating extreme racial disparities in the prison system. Although Northpointe issued a rebuttal to the ProPublica study in which the company 'unequivocally rejects the ProPublica conclusion of racial bias in the COMPAS risk scales', there is still the underlying challenge to independent auditing.

With COMPAS, again, one of the obstacles to challenging learned bias and ensuring all defendants' equal due process rights is that Northpointe's algorithm is proprietary and not open to independent auditing. And while judges are often presented with the COMPAS readouts during hearings, this material is not always shared in full with the defendants or their counsel, which provided the grounds for the ultimately unsuccessful appeal in *Loomis v. Wisconsin* to the US Supreme Court in 2017. The ACLU and over one hundred other organizations in the US have called for an end to such pre-trial risk assessment tools.

These types of risk assessment algorithm are not only being deployed for domestic criminal justice systems. Since 2013 the US Immigration and Customs Enforcement (ICE) has relied on such tools to make immigration detention decisions. Shockingly, following President Donald Trump's nationalist stance on immigration, ICE has since changed its algorithm to now always recommend detention, regardless of an individual's criminal history. This, in part, has contributed to the massive spike in immigration detention and human rights abuses at the US border. This is a reminder that the parameters of machine-learning algorithms themselves can easily be adjusted for political and discriminatory means against minority populations.

Another challenge is how data is collected and presented. Across Europe, anti-immigration populist movements and governments cite hundreds of thousands of migrants entering the EU, or the million or more asylum applicants each year, to stoke anti-immigration fears that lead to violence against minorities and the passage of restrictive laws or policies, such as iBorderCtrl. But in 2017 researchers in the UK noted the flaw in how such data is being generated and broadcast. Frontex, the EU's border security agency, can count the same person multiple times. For example, the migrant or refugee who arrived in the EU at Greece and left it to look for work in Albania, only to return through Croatia or Hungary, may be counted as two or more people entering the EU. Similarly, the presentation of asylum data is a reflection of the total number of applications across the EU and not the total number of individuals, and many asylum seekers may register in multiple countries. In these examples, the data used to inform machine-learning algorithms at borders or used in political campaigns or legislation can be flawed, and in an environment of structural bias against minorities such misrepresentation of data can fuel disinformation, hate speech and violence.

When big data is drawn from existing systems of ethnic, gender or other inequalities the bias is replicated: bias in, bias out.

The dangers of 'big data'

The challenges around surveillance and discriminatory algorithms are underpinned by the increasing availability of 'big data' – not only from official records and coercive intrusions by governments, but also indirectly through the surreptitious collation of microdata on issues such as travel patterns, smartphone usage and the like. This is an area where private corporations, rather than states themselves, often play a leading role, and are developing 'products' that may conceal agendas that have profound implications for human rights.
One area where these challenges of big data are on full display is the 'smart city', deemed smarter because it relies on an expanding network of interconnected devices, sensors and scanners to gather data on individuals and their environment, to adjust or report according to the relevant protocol. This is part of the Internet of Things, but for all its utopian ideals of maximizing environmental sustainability it can also produce a dystopian surveillance nightmare, as in Xinjiang. And as such, as the tech industry seeks to combine technology with urban planning, its pursuit of innovation appears to outpace solutions for privacy and other rights concerns.

Israel, a leading technology hub and world producer of surveillance tools, is also increasingly turning to smart city design in Jerusalem that, as digital rights activists point out, increasingly reaffirm inequalities between Israeli citizens afforded privacy rights and due process and West Bank Palestinians who have few such rights. Meanwhile in Canada, Google's sister firm Sidewalk Labs has been developing Waterfront Toronto as a fully data-fuelled smart neighbourhood, but concerns over its human rights impact sparked the #BlockSidewalk movement. Canadian author and digital rights activist Cory Doctorow described it as a 'terrible idea to let vast, opaque multinational corporations privatize huge swathes of our city, webbing them with surveillance sensors and subjecting them to opaque, unaccountable algorithmic analysis and interventions'.[16] In May 2020, Sidewalk Labs scrapped the project due to the economic uncertainty in the wake of the Covid-19 pandemic. While the context in Toronto may seem very different to Jerusalem, there are still concerns around the implications of surveillance and discrimination: as highlighted by one commentator, after the cancellation of the project was announced, 'minority groups and people of colour face more threats from surveillance than majority groups, and a digital stop-and-frisk program could subject some people to more oversight than others'.[17]

In India, a Smart Cities Mission was launched in 2015 with plans to 'modernize' 100 cities by 2020, but the lack of consideration for all residents in the plans, especially for

[16] #BlockSidewalk website. Available at https://www.blocksidewalk.ca/supporters

[17] Lachman, R., 'Sidewalk Labs' city-of-the-future in Toronto was a stress test we needed', Policy Options, 28 May 2020.

A man's face glows as he goes through a biometric turnstile on his way to Jerusalem at Qalandiya checkpoint, Palestine.

Eddie Gerald

already marginalized Dalits, Adivasis and religious minorities, demonstrates that 'smart' does not necessarily mean 'more equal.' As India's Housing and Land Rights Network (HLRN) noted in a 2018 report:

'With one in six urban Indians still living without adequate housing and access to essential services, and high rates of violence and crime being reported against women and children, especially belonging to Dalits or other minorities, in urban areas, a "smart city" cannot just be about installing seamless digital connectivity, or making physical infrastructure more efficient and reliable.'

In sensible advice for any would-be smart city planners around the world, HRLN cautions: 'When marginalized individuals, groups and communities are not at the centre of any scheme, it is unlikely that it will address their concerns and achieve inclusion and an improved quality of life, as claimed in the Smart Cities Mission's objectives.'[18]

While South Korea's Songdo International Business District, a smart city built on reclaimed land from the Yellow Sea, may not avoid some of the concerns noted above, South Korea does offer a useful framework for would-be smart city developers. The country hosts the annual World Human Rights Cities Forum, which adopted the Gwangju Guiding Principles for a Human Rights City in 2014. The Gwangju Principles reaffirm the need to respect the principle of equality and equity among all residents, implement non-discrimination measures including gender-sensitive policies and protection for minorities and vulnerable groups, with human rights mainstreamed into all aspects of planning, implementation and monitoring. In other words, as technologies and big data create new tools, rather than merely embracing

[18] Housing and Land Rights Network, *India's Smart Cities Mission: Smart for Whom? Cities for Whom?* (update 2018), Housing and Land Rights Network, 2018.

A 2017 study by Brookings found the income penalty for minority STEM PhDs taking on university employment in the US tends to be **US$13,000 more a year** than for non-minority STEM PhDs.

digitization to make cities smarter we should be embracing these tools to make them Human Rights Cities. Online, big data and algorithmic bias is also a problem. In 2015, it was revealed that the Google Photos algorithm had labelled two black friends as gorillas. The company was quick to apologize, but the root of the problem remains across multiple tools where intersectional bias is even more pronounced and many online facial recognition algorithms are far more likely to falsely identify or match black women. The reason, it has been argued, is that 'the values of the web reflect its builders – mostly white, Western men – and do not represent minorities and women'.[19] This has a similar cause to the example of automated recruitment algorithms noted above, when big data is drawn from existing systems of ethnic, gender or other inequalities the bias is replicated: bias in, bias out.

Big data is the driving force behind the growth of AI, and because it is increasingly affecting everyone's lives, says Adrian Weller of the UK's Alan Turing Institute, 'it is very important that we have a diverse set of stakeholders designing and building them'.[20] Unfortunately, as noted in a 2019 study by the AI Now Institute, 'there is a diversity crisis in the AI sector across gender and race', with no public data even available for trans or other gender minorities.[21] This lack of diversity is common across the whole science, technology, engineering and mathematics (STEM) field in general, but even more so at universities where the lack of diversity in STEM faculties can arguably be said to impact minority students choosing the field as a career path. A 2017 study by Brookings found one startling revelation: the income penalty for minority STEM PhDs taking on university employment in the US (rather than entering the private sector) tends to be US$13,000 more a year than for non-minority STEM PhDs.[22] But this is only part of the issue. As presented above, bias can be intersectional and certainly one way of addressing the replication of this bias is to ensure more intersectional diversity in the big data workforce.

In China, the situation is worse. Uyghurs are largely prohibited from

[19] Snow, J., 'Bias already exists in search engine results, and it's only going to get worse', MIT Technology Review, 26 February 2018.

[20] Ram, A., 'AI risks replicating tech's ethnic minority bias across business', *Financial Times*, 31 May 2018.

[21] West, S.M., Whittaker, M. and Crawford, K., *Discriminating Systems: Gender, Race and Power in AI,* AI Now Institute, 2019, p. 3.

[22] Startz, D., 'Why is minority representation lagging among STEM faculty? It could be the money', Brookings, 15 December 2017.

We must protect against technology development creating new dependencies and inequalities, not only in terms of the 'digital divide' – put simply, the separation between the haves and have-nots of certain technologies – but also the more nuanced issue of 'digital colonialism'.

even enrolment in STEM programs. This discrimination is part of China's overall essentializing of ethnic and religious minorities, whereby their career and cultural place is relegated often to merely one of entertainment and food. While China proclaims its interest in becoming a world leader in advanced technologies, the denial of STEM education opportunities for Uyghurs guarantees their marginalization from any residual economic benefits that might be associated with even relatively innocuous technologies. Instead, Uyghurs have in fact been the principal surveillance target of many of these technologies. For these reasons, Uyghur students who wish to pursue academic studies in engineering or aerospace, for example, must seek opportunities abroad, such as in Turkey, but this also introduces a vicious cycle of repression: having a family member studying abroad has become reason enough to interrogate or detain Uyghurs in China.

Another problem is that as the industry expands to create new well-paying jobs, this lack of diversity reaffirms historical economic inequalities of employment sectors that already reinforce gender stereotypes and whose workers are predominantly drawn from minority communities. It becomes a vicious cycle, bad data feeding algorithms that shape real-world experiences, generating new bad data, and so on. In addition to greater diversity in the workforce, legislation is needed to address algorithmic bias. In early 2019, the US state of Washington, home to companies like Amazon and Microsoft, introduced an algorithmic accountability bill that would establish guidelines for the procurement and use of automated decision-making systems. The lawmakers recognized the risks to 'due process, fairness, accountability and transparency, as well as other civil rights and liberties'. A major provision of the bill would ensure that such tools employed by the public sector, such as pre-trial risk assessment programs in the criminal justice system, would be available before, during and after deployment for third-party auditing and research. Following such state-led legislative agendas, the US Congress has introduced the federal-level Algorithmic Accountability Act, which, if adopted, would task the Federal Trade Commission with the creation of rules for evaluating algorithms for bias or discrimination, including the datasets used to train machine learning.

Meanwhile, across Europe, many courts are finding that the human rights impacts of unchecked big data outweigh any potential benefits to the government. In February 2020, for example, a Dutch court in *NJCM v the Netherlands* shut down the country's System Risk Indication (SyRI) system, which had relied on big data to predict benefit fraud. Many of its targets had been ethnic and religious minority Dutch citizens who are more often among the poor and vulnerable groups of society targeted by such automated welfare systems. In 2019, Swedish and French data protection authorities fined and halted programs involving facial recognition systems to gather and process biometric data about student attendance. Such victories for the right to privacy in Europe are made possible by the General Data Protection Regulation (GDPR), which covers, among other things, an individual's right to receive information about the kind of data collected about them and how it will be used. While the GDPR is still in its infancy, along with the European Parliament's work in formulating a framework for algorithmic accountability,[23] Europe is leading the charge in addressing many of the concerns of big data examined in this chapter and setting standards that can hopefully provide models for protecting vulnerable minority populations elsewhere.

Reversing the trend: how technologies can be used to defend human rights

This chapter has profiled a number of concerning trends at the intersection of technology and human rights, with particularly troubling implications for minorities and indigenous peoples. These are serious issues that require considerable research, legislation and tools to combat and remedy them. At the same time, many of these technologies are offering new connectivity, platforms and resources to improve livelihoods and rights defence for many. But if these new tools and technologies are to be developed or repurposed for these objectives, then minorities and indigenous peoples must be informed and involved at every step, from design to implementation and evaluation. We must protect against technology development creating new dependencies and inequalities, not only in terms of the 'digital divide' – put simply, the separation between the haves and have-nots of certain technologies – but also the more nuanced issue of 'digital colonialism'. The latter raises a range of concerns – bound up in the technologies themselves, not simply their lack of availability – around power inequities, discrimination and the marginalization of non-majority voices. For instance, it is not enough to provide universal access to the internet; it is also necessary to ensure that the online world is safe, accessible and non-discriminatory for minorities, indigenous peoples and other groups.

An example of how this can be achieved is the development, since 2016 of a mobile application called #thismymob, by researchers at the faculty of Engineering and Information Technology at the University of Technology Sydney. The project, explains its director Christopher Lawrence, was born from the concept of 'postcolonial computing', and uses participatory design to create new digital technologies with and for indigenous peoples. Participatory design, explains Lawrence, 'ensures that the technology we design is culturally appropriate, and usable in

[23] European Parliament Research Service, *A Governance Framework for Algorithmic Accountability and Transparency*, European Parliament, 2019.

VICTORIES FOR THE RIGHT TO PRIVACY IN EUROPE

Victories for the right to privacy in Europe made possible by the General Data Protection Regulation (GDPR):

- In February 2020, in NJCM v the Netherlands, a Dutch court shut down the country's System Risk Indication (SyRI) system, which had relied on big data to predict benefit fraud. Many of its targets had been ethnic and religious minority Dutch citizens who are more likely to be among the poor and vulnerable groups of society targeted by such automated welfare systems.
- In 2019, Swedish and French data protection authorities halted and fined programs involving facial recognition systems to gather and process biometric data about student attendance.

a wide variety of communities and contexts…. We recognize that having indigenous leadership on research and development projects is fundamentally important.' The platform, which was developed taking intersectional identities of region or gender into account, allows indigenous users to connect with elders around the country for guidance and support – for example, encouraging indigenous students to pursue careers in STEM or facilitating artists to promote their work to both indigenous and non-indigenous communities.[24]

The Human Rights Investigations Lab at the University of California Berkeley engages in multidisciplinary practicums to prepare students to mine social media for documentation of human rights violations. The lab has partnered with leading international human rights organizations and media. For example, an explosive 2018 report by Reuters on hate speech in Myanmar was based on efforts by the lab to collect and translate over 1,000 social media posts involving hate speech against Rohingya. They have also overseen investigations into Sudan, Syria and elsewhere. Much of the lab's research is based on open source material, and the lab is also working to develop an international protocol on open source investigations. Its methods can be employed by anyone, and by demystifying and disseminating such skills beyond the university setting it creates a toolkit for minority and indigenous activists to increasingly employ technology themselves in their rights defence. Human rights organizations like WITNESS have also developed new tools and training for rights defenders to better document and disseminate human rights concerns on social media.

Researchers are working on how machine learning, too, could be exploited for positive human rights outcomes – for example, by developing algorithms to process large amounts of social media or video content in order to flag hate speech or evidence of human rights abuses. Blockchain, perhaps better known as the technology behind cryptocurrency (which has also attracted criticism for its potential use in illicit transactions), allows for the establishment of

[24] Lawrence, C., '"Digital land rights": co-designing technologies with Indigenous Australians', *The Conversation,* 31 July 2018.

A Rohingya woman living in a refugee camp in Cox's Bazar, Bangladesh.

Credit: Ramazan Nacar

anonymized, secure and decentralized networks, and also has human rights applications. For example, video evidence or other social media content that reveals human rights abuses against minority or indigenous populations could be verified and entered into dedicated blockchain networks, creating decentralized, open source, tamper-proof pools of big data, potentially useful for anything from advocacy to international litigation.

AI is also being developed for human rights applications. In 2016, AI research at the University of Sheffield in the UK and University of Pennsylvania in the US trained an algorithm on trial data from the European Court of Human Rights to predict judicial decisions with 79 per cent accuracy. Rather than falling victim to some of the concerns of machine learning noted above, such algorithms at regional or national courts could be used to help human rights lawyers better prepare their cases before submission and increase the effectiveness of human rights litigation.

Another example of the innovative use and increasing ease of access to technologies once reserved for governments and militaries is the benefit of satellite imagery in documenting the scale of mass internment in Xinjiang. Throughout 2018 in particular, Shawn Zhang, a graduate student at the University of British Columbia law school, relied on open source satellite imagery to document multiple large-scale internment camps across Xinjiang at a time when the Chinese government was still categorically denying their existence. The research of scholars like Zhang or human rights organizations such as Fortify Rights and HRW, which have also used satellite images in documenting the forced displacement of Rohingya in Myanmar, demonstrates how technology can provide unequivocal evidence of gross violations against minority or indigenous populations even in areas where the majority government refuses independent access and fact-finding. Such data is valuable for human rights documentation as well as later accountability and transitional justice mechanisms.

Meanwhile, digital security remains at the frontline of risk and protection for minority and indigenous rights defenders and their allies. And here, again, the risk analysis and design of new tools must be conducted with the full consent and participation of minority and indigenous stakeholders. End-to-end encryption, for example, should be a fundamental right because in a digital age it is one of the bulwarks against infringement of the freedom of expression, association, assembly, and the right to privacy, not to mention real-world ramifications. Meanwhile, even the best encryption or strongest passphrase is ultimately meaningless if a computer or mobile device is compromised, as seen above with the NSO Group hacking of WhatsApp. Additionally, police and state agents in many regimes seldom hesitate to use physical force such as torture or threatening one's family members to extract information including passphrases. Digital security without physical or psychological security is not enough, and this has given rise to the concept of holistic security. This is one area of digital rights and security, among many, that is still at risk and remains crucial to the protection of minority rights. Groups such as the Guardian Project, EFF, Tactical Tech Collective and others continue to work with frontline rights defenders to develop new tools and holistic security routines, adapting to the digital age.

Conclusion

What these examples demonstrate is that, while some technologies may raise particular concerns, first and foremost it is the governance and protections around them that are likely to impact most directly on minorities and indigenous peoples, for better or worse. This is illustrated by the challenges around monitoring negative content about minorities and indigenous peoples online. While the dangers of hate speech and misinformation are very real, contributing to the continued exclusion of many communities and even to physical violence against them, restricting freedom of expression and undermining privacy rights in the name of preventing hate speech – a tactic employed by many authoritarian governments to justify internet shutdowns and other draconian policies – is no solution. Indeed, more often than not, such measures serve only to further silence and disenfranchise the groups most at risk.

Since many of the technologies discussed in this chapter are new and constantly evolving, further research, documentation and the formulation of dedicated guidelines to ensure minority and indigenous rights within their design and implementation will be of long-term benefit. It should not be assumed, fatalistically, that protection regimes will never be capable of catching up with technological developments. In fact, international human rights law is already highly capable of guiding these technologies and protecting minorities and indigenous peoples in the digital age. If even existing human rights law were better applied, we might find the need for new rules and guidelines were largely redundant.

Though many governments have clearly been directly complicit in using technology to perpetrate human rights abuses against minorities and indigenous peoples, an added issue here is that technological design, development and roll-out involves an increasing array of non-governmental actors, including corporations and research institutions. These independent organizations are often vested with considerable powers to guide decision-making in areas such as law enforcement, migration management and welfare provision – traditionally the preserve of governments – without many of the oversight, accountability or regulatory protocols that would be applied to public bodies as a matter of course. What is needed, then, is far better transparency at every stage, not only in the design and implementation of these products, but also in how companies make their rules for oversight and decision-making. This means, for example, that companies need to disclose when and how they work with governments, as well as what information they collect and share. In addition, there must be means for effective challenge or remedy for these decisions.

Technology alone, however well-designed, will not address underlying societal injustices, and in many cases may in fact perpetuate or worsen inequalities for minority and indigenous communities. Just as human rights law should govern the design and implementation of new technologies, so too should it govern broader social norms and the increasing integration of technologies into our lives. Some basic principles to support this process include:

- *Uphold freedom of expression and information as a 'default setting' for the use of any technologies.* International law is unequivocal that freedom of expression and information can be restricted only under the most extreme circumstances, and that any restrictions should be prescribed by law, pursue a legitimate aim, and be necessary and proportionate. However, many government efforts to regulate speech online have failed to meet these standards. In particular, the increasing use of internet shutdowns by authorities to quell dissent should be seen not only as vehicles for human rights violations, but as violations in and of themselves.

- *Ensure that the highest standards of corporate responsibility are imposed on those working in areas of technology with potential human rights impacts.* In particular, the use of private-public partnerships for predictive policing or surveillance-based security systems should not enable governments to outsource their human rights responsibilities to opaque and unaccountable institutions. Private companies working in sectors with potential impacts on human rights protections should be held to the highest standards on issues such as transparency, due diligence and public regulation.

- *Streamline human rights law more effectively into the development, use and delivery of new technologies.* While the evolving nature of some emerging technologies may require new legislation and frameworks, it is important to recognize that there is a wealth of existing human rights law that, if effectively implemented, could support the realization of a more inclusive and socially beneficial future. For instance, in the case of private actors, the UN Guiding Principles on Business and Human Rights call on businesses to prevent and mitigate the actual and potential human rights abuses associated with their business practices, and to conduct regular, effective, independent human rights impact assessments of all their operations. This is increasingly necessary for technology companies and internet providers.

- *Impose clear guidelines on the ethical, non-discriminatory use of personal data by companies, governments and other actors.* While the historic lack of disaggregated data for minorities and indigenous peoples has been a major barrier to their efforts to secure adequate political representation, public spending and other rights, it is important that the opportunities offered by the latest 'smart' data-collection tools are used in a rights-based framework that respects privacy and non-discrimination. The potential opportunities of AI and big data to increase visibility should not be undermined by excluding individuals or communities from particular benefits or services through the use of discriminatory or biased algorithms.

- *Establish clear principles of accountability for any decision-making assisted by AI, algorithms and other technologies to ensure that the rule of law is upheld.* In particular, any negative decisions involving predictive policing, parole and immigration that lead to continued incarceration, visa rejections, deportation or detention should be followed up by an appeals process overseen by a human adjudicator.