Business and Human Rights in the Digital Environment

Accompanying notes

February 2020

Contents

Introduction 03

Glossary 04

Module 1: Introduction to Business and Human Rights in the Digital Environment 06

- International human rights law 06
- The application of international human rights law to businesses: a history 06
- The UN Guiding Principles 10
- Proposals for a treaty 10
- BHR in the digital environment: a recent history 10
- Emerging issues 13

Module 2: Privacy and Free Expression in the Digital Age 15

- Privacy in the digital age 15
- Freedom of expression in the digital age 31

Module 3: The UN Guiding Principles Pillar I: the State Duty to Protect Human Rights 47

- Foundational principles 47
- Operational principles 48
- National Action Plans on BHR 51
- Timeline of National Action Plans 54
- Attempts to establish a state duty to protect against human rights abuses by businesses 57

Module 4: The UN Guiding Principles Pillar II: the Corporate Responsibility to Respect Human Rights **60**

Foundational principles 60

Operational principles 63

Module 5: The UN Guiding Principles Pillar III: Access to Remedy **70**

Foundational principles 71

Operational principles 73

Introduction

Businesses can impact human rights wherever and however they operate. Whether it is through the products and services they provide or their own internal policies and processes, companies can impact the entire range of human rights positively or negatively.

In recent years, there have been a number of high-profile examples of businesses in the tech sector impacting human rights, with effects often being felt far beyond the countries in which they are based. These impacts have had a particular effect on the rights to privacy and freedom of expression, and include large-scale data breaches, the sharing of personal data without consent, censorship by online platforms, and companies assisting governments in undertaking online surveillance. All of these examples demonstrate the significant and widespread impact that tech companies' actions can have on the enjoyment of human rights.

Historically, however, the international human rights system paid little attention to the actions of businesses, largely because businesses, unlike states, do not have legal obligations under international human rights law. Since the 1970s, however, there has been an increased interest in the impact of businesses on human rights, and it is now increasingly accepted that businesses have a responsibility to respect human rights, including by businesses themselves. A key milestone in this journey was the development of the United Nations Guiding Principles on Business and Human Rights (UNGPs) in 2011, and their endorsement by the UN Human Rights Council. While not binding, the UNGPs set out the duties of states and responsibilities of businesses when it comes to respecting, protecting and fulfilling human rights.

The modules and slides on Business and Human Rights in the Digital Environment, taken alongside these accompanying notes, provide a comprehensive overview of the human rights framework relevant to businesses, and specifically the tech sector, helping to inform the reader's understanding and engagement.

Glossary

AI	Artificial intelligence
FIPPs	Fair Information Practice Principles
GDPR	General Data Protection Regulation
GNI	Global Network Initiative
ICCPR	International Covenant on Civil and Political Rights
ICESCR	International Covenant on Economic, Social and Cultural Rights
IoT	Internet of Things
OECD	Organisation for Economic Co-operation and Development
MLAT	Mutual Legal Assistance Treaty
NAP	National Action Plan on Business and Human Rights
TNC	Transnational corporation
UDHR	Universal Declaration of Human Rights
UN	United Nations
UNGPs	United Nations Guiding Principles on Business and Human Rights
UN HCHR	United Nations High Commissioner for Human Rights
UNSR on BHR	United Nations Special Representative on Business and Human Rights

Module 1: Introduction to Business and Human Rights in the Digital Environment

Module 1: Introduction to Business and Human Rights in the Digital Environment

The business and human rights framework has developed as part of the broader international human rights framework. As such, a basic understanding of international human rights law, and its relevance to businesses, is essential to understanding the roles and responsibilities of states and businesses as set out in the UN Guiding Principles on Business and Human Rights (UNGPs). After a brief review of international human rights law, this module looks at the history of the business and human rights framework leading up to and including the development of the UNGPs themselves.

International human rights law

International human rights law is a field of international law which sets out the human rights to which all persons are entitled, and which states should respect, protect and fulfil. These human rights are set out, for the most part, in treaties, such as the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR). These treaties are ratified by states, and are then binding upon those that have ratified them. The obligations largely focus on prohibiting the states from violating the human rights of individuals in their territory and subject to their jurisdiction. However, those obligations are only binding upon states and not other parties (such as businesses or individuals).

The ICCPR and the ICESCR were both drafted in the 1960s, and drew inspiration from an earlier UN non-binding document, the Universal Declaration of Human Rights (UDHR) which had been adopted in 1948. Since the ICCPR and the ICESCR, other "core" human rights treaties have been drafted, and there are now a total of nine.[1]

[1] These are the International Convention on the Elimination of All Forms of Racial Discrimination (1965), the International Covenant on Civil and Political Rights (1966), the International Covenant on Economic, Social and Cultural Rights (1966), the Convention on the Elimination of All Forms of Discrimination against Women (1979), the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (1984), the Convention on the Rights of the Child (1989), the International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families (1990), the International Convention for the Protection of All Persons from Enforced Disappearance (2006) and the Convention on the Rights of Persons with Disabilities (2006).

The application of international human rights law to businesses: A history

Historically, since treaties only bound states, international human rights law was only considered as relevant when it came to states and the actions of states in particular. In the decades since the earliest human rights treaties, an acceptance emerged that the obligations of states to ensure the protection of human rights also included ensuring that individuals' human rights are not violated by *other* parties, such as businesses or other individuals. With this acceptance arose questions over whether these third parties - and particularly businesses - had some form of duty, responsibility or role when it came to human rights, themselves, distinct from that of states, and if so, how to articulate it.

UN Commission on Transnational Corporations

Consideration of if and how the international human rights framework applied to the activities of businesses became an area of interest in the 1970s following a series of high-profile events involving transnational corporations (TNCs) and accusations of exploitation. In 1973, the UN Economic and Social Council appointed a Group of Eminent Persons tasked to study the impact of TNCs on economic development and international relations, and to provide advice to the UN. One of their recommendations was that the UN should establish a permanent Commission on TNCs (supported by a Centre) which would provide support to the UN and states on matters involving TNCs and foreign investment, including through considering the development of a multilateral agreement on TNCs. In 1974, both the Commission and Centre on TNCs were established.

At its first session in 1975, the Commission on TNCs decided to prioritise the development of a Code of Conduct on Transnational Corporations, and the following year it established an Intergovernmental Working Group to develop the Code. The Group's negotiations started in 1977 and, in 1982, it developed a first draft Code for the consideration of the governments in the Commission on TNCs. That draft Code, and further revisions, were periodically discussed during special sessions of the Commission over the following years. However, due to significant opposition to the draft Code - particularly by states in the Global North - it was abandoned in 1994 and the Commission and Centre on TNCs dismantled.[2]

Global Compact

Under the leadership of Secretary-General Kofi Annan, the UN embarked on a new effort to address the societal impacts of the corporate sector, which resulted in the

[2] For more information, see Sauvant, K., "The Negotiations of the United Nations Code of Conduct on Transnational Corporations: Experience and Lessons Learned", *The Journal of World Investment & Trade* 16 (2015) pp. 11-87, available at: http://ccsi.columbia.edu/files/2015/03/KPS-UN-Code-proof-2-Journal-of-World-Investment-and-Trade-March-2015.pdf.

development of the UN Global Compact in 2000. The Global Compact sets out ten non-binding principles intended to guide businesses in the development of socially and environmentally sustainable practices. The Global Compact includes two principles on human rights, which state that businesses should support and respect the protection of internationally proclaimed human rights (Principle 1) and make sure that they are not complicit in human rights abuses (Principle 2).

Although the Global Compact was relatively well-received by the business community, and a significant number of companies joined the initiative, some human rights defenders opposed the voluntary nature of the Global Compact, arguing that a binding instrument was necessary to stem harmful corporate activity.

Draft Norms

Following criticism of the voluntary nature of of the previous initiatives dealing with corporate accountability for human rights impacts, the UN Sub-Commission on the Promotion and Protection of Human Rights (the UN Sub-Commission) turned to the issue. (The UN Sub-Commission was the main subsidiary body of the former Commission on Human Rights, replaced in 2006 with the Human Rights Council). In the early 2000s, the UN Sub-Commission drafted the Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises with Regard to Human Rights (UN Draft Norms). Although the UN Draft Norms did not have a binding effect, and were scuppered by a resolution from the UN Commission on Human Rights, they represented the first attempt by the UN to identify potentially binding human rights standards that could apply to corporations.

Special Representative on Business and Human Rights

Following the failure of the UN Draft Norms, the Commission on Human Rights requested the UN Secretary-General to appoint a special representative on the issue of human rights and transnational corporations and other business enterprises. In 2005, Kofi Annan did so, and appointed John Ruggie as the UN Special Representative on Business and Human Rights (UNSR on BHR).

The initial mandate of the UNSR on BHR was relatively limited in scope, focusing largely on the identification and clarification of standards of corporate accountability with respect to human rights and the identification of state and corporate best practices.

During his first term, Ruggie developed the Protect, Respect, and Remedy Framework, which set out a state duty to protect against human rights abuses, a corporate responsibility to respect human rights, and the need for victims' access to judicial and non-judicial remedies.

After his mandate was extended, Ruggie operationalised the Framework by developing what are now known as the UN Guiding Principles on Business and Human Rights (UNGPs). They are a set of non-binding standards based on the Framework and were endorsed by the UN Human Rights Council in 2011.[3]

The application of international human rights law to businesses: the United Nations Guiding Principles

The UNGPs are a framework that affirms that, just as states have a duty to *protect* human rights, companies also have a responsibility to *respect* human rights. The principles are grounded in the acknowledgement that the actions of business enterprises can significantly impact human rights. The UNGPs articulate the obligations and responsibilities of government and businesses with regards to business and human rights, and lay out operational principles for governments and businesses to meet them. They comprise three pillars:

- A state duty to protect against human rights abuses by third parties, (including businesses), through appropriate policies, regulation, and adjudication;
- An independent corporate responsibility to respect human rights, which means that business enterprises should avoid infringing on the rights of others and address adverse impacts with which they are involved;
- Access for victims to effective remedy, both judicial and non-judicial.

We will look at each of these pillars in detail in later modules. For now, it is important to understand that:

- The UNGPs are comprised of thirty one principles, each with commentary elaborating its meaning and implications for law, policy, and practice.
- They encompass all internationally recognised rights, and apply to all states and all business enterprises, of whatever size.
- The UNGPs do not, by themselves, create new legally binding obligations for businesses, however they are influential and set out accepted expectations by states and other key actors, including businesses themselves.

The UNGPs have not gone without criticism, particularly among civil society organisations. Many organisations have lamented the voluntary nature of the UNGPs, and would prefer to see a binding instrument. There has also been criticism of the lack of implementation of the UNGPs by many states and businesses, leading

[3] UN Human Rights Council, Resolution 17/4, Human rights and transnational corporations and other business enterprises, UN Doc. A/HRC/RES/17/4, 6 July 2011.

to impunity or minimal change, again as a result of the lack of any means of enforcement of the UNGPs.

The application of international human rights law to businesses: proposals for a treaty

In September 2014, Ecuador proposed that the UN Human Rights Council establish an open-ended intergovernmental working group "to elaborate an international legally binding instrument to regulate, in international human rights law, the activities of transnational corporations and other business enterprises".[4] Bolivia, Cuba, South Africa and Venezuela also supported the proposal, and it was carried by a majority of twenty states in support, fourteen opposed, and thirteen abstaining. Support came largely from Africa, China, India and Russia, whereas opposition was found among the European states, the United States, Japan and South Korea. Those states in favour tend to argue that a non-binding instrument is insufficient, and are often states where transnational corporations from the Global North have operations, meaning they have a particular interest in ensuring that any harms stemming from those operations can be addressed. Those against tend to argue that a treaty is not necessary and that more can be done using the existing voluntary frameworks; often states who are against a new treaty - particularly the United States - are the home states of many transnational corporations, and so would be particularly affected by new binding obligations relating to their operations. By 2018, the group had developed a "zero draft" of a legally binding instrument, which was revised in 2019.

Business and human rights in the digital environment: A recent history

While the internet and digital technology have had impacts upon human rights since their earliest days, it is only in recent years that those impacts have really caught the attention of human rights defenders. Two specific human rights - the rights to privacy and freedom of expression - have been particularly affected by the digital environment, and later modules will look at these rights more closely. In the next slides, however, we examine some of the broader issues and developments in the field of business and human rights in the digital environment over the last decade.

Today, there are over 5 billion unique mobile phone users in the world, and around half of the world's population uses the internet, figures which have increased dramatically over the last 20 years. With such growth in the use of the internet and

[4] UN Human Rights Council, Resolution 26/9 Elaboration of an international legally binding instrument on transnational corporations and other business enterprises with respect to human rights, UN Doc. A/HRC/RES/26/9, 14 July 2014.

digital technology, their impacts upon the rights to privacy and freedom of expression have become pronounced. These rights are particularly significant for a number of reasons, including:

- Our online experience relies upon the generation, collection, processing and sharing of large amounts of data - often personal and sensitive data - that can be used to profile individuals or groups, engaging our right to privacy in ways never before seen.
- The rights to privacy and freedom of expression are "gateway rights", which enable the exercise and enjoyment of many more. Making it easier for people to communicate, particularly privately, has created opportunities for individuals to exercise their rights to freedom of association, and peaceful assembly, for example.

2011: It was in 2011, during the Arab Spring, that the potential of the internet and digital technology was harnessed on a scale not seen before. Tens, if not hundreds, of thousands of people across the region used their devices to organise, assemble, campaign and protest. Various tools were developed and used as part of these efforts, such as encrypted communications tools. However, some tech companies cooperated with governments, providing surveillance technology to help identify protestors or information on users and their communications.

2012: The year 2012 saw various efforts by governments to control the free flow of information online and, in particular, to control the rise in user-generated content shared on online platforms. A number of states developed legislation which either enhanced levels of online surveillance, or restricted online anonymity. The term "deep packet inspection" was increasingly used, particularly in reference to the UK government's proposals to surveil the internet. At the UN, the Human Rights Council passed a resolution on human rights and the internet which affirmed that "the same rights that people have offline must also be protected online".[5]

2013: 2013 was the year in which Edward Snowden, a whistleblower who leaked classified information from the US National Security Agency when he was a Central Intelligence Agency employee and subcontractor, revealed that a number of intelligence agencies had engaged in mass data gathering practices. The revelations raised questions about the commitment of governments to the protection of the right to privacy and the potential chilling effect on freedom of expression and freedom of association.

More tech companies joined the handful that pioneered the practice of releasing transparency reports, which publish the number of government requests or judicial

[5] UN Human Rights Council, Resolution 20/8, The promotion, protection and enjoyment of human rights on the Internet, UN Doc. A/HRC/RES/20/8, 16 July 2012.

orders to take down or block content under local legislation or obtain access to and monitor user data that the company has received and complied with. At the UN, a General Assembly resolution to strengthen the right to privacy in the digital age, co-sponsored by Germany and Brazil, received widespread support.

2014: Concerns continued to rise during 2014 over mass surveillance practices of government intelligence agencies. In a departure from the past, when most governments preferred a behind-the-scenes approach to internet control, countries began to adopt new laws that legitimised existing repression and effectively criminalised online dissent. 2014 also saw increased government pressure on independent news websites, which had previously been among the few uninhibited sources of information in many countries, in addition to more people detained or prosecuted for their digital activities than ever before.

A 2014 report by the UN Office of the High Commissioner for Human Rights on The Right to Privacy in the Digital Age stated that mass surveillance was "emerging as a dangerous habit rather than an exceptional measure" and that practices in many states revealed "a lack of adequate national legislation and/or enforcement, weak procedural safeguards, and ineffective oversight".[6]

2015: Over 2015, online content removal requests by governments increased, arrests and intimidation escalated, surveillance laws and technologies multiplied, and governments undermined encryption and anonymity. 2015 was, however, also the first year of Ranking Digital Rights Corporate Accountability Index. The Index concluded that "there were no winners", as even companies in the lead were falling short, and argued that companies needed to improve their commitments to, and disclosures of, policies and practices that affect users' freedom of expression and privacy. Also, in 2015 the increase in usage of the terms "big data", where large datasets are analysed to provide insights into a particular topic, and "Internet of things", where an increasing number of devices and appliances are connected to the internet, raised new questions about privacy and other human rights, particularly on how to ensure effective privacy safeguards on the collection, storage, use and sharing of personal data.

2016: This year, governments increasingly focused their attention on messaging apps like WhatsApp and Telegram, which spread information quickly and securely. Online manipulation and disinformation tactics played an important role in elections in several countries. There was an increase in focus on hate speech targeting women and minorities from governments and civil society groups. Women were targeted with misogynist abuse and hate speech, and prominent women stepped away from some popular social networking sites. Finally, in 2016 it

[6] Office of the United Nations High Commissioner for Human Rights, "The right to privacy in the digital age", UN Doc. A/HRC/27/37, 30 June 2014.

became evident that political parties and movements were using big data to analyse information and individual choices to sharpen and tailor political messages to influence electoral outcomes, as seen in 2016 in the UK's Brexit referendum and the US presidential election.

2017: In 2017, there was a rise in disruptions to mobile internet service. And while propaganda, surveillance, and restricting access to information were not new phenomena, big data and technology relying on algorithms allowed companies and governments to process vast amounts of data efficiently and quickly, and to take more nuanced actions targeting individuals.

2018: Ranking Digital Rights' 2018 Corporate Accountability Index concluded that more than half of the companies evaluated improved disclosure in multiple areas affecting users' freedom of expression and privacy, yet it also argued that companies are not transparent enough about the design, management, and governance of digital platforms and services that affect human rights. The UN Special Rapporteur on the Promotion and Protection of the right to freedom of opinion and expression published a report that focuses on establishing a human rights approach to platform content regulation.[7]

Emerging issues

The term "new technologies" originated in the 1950s, and refers to technology that radically alters the way something is produced or performed, especially by labour-saving automation or computerisation. "Emerging technologies", a similar term, refers to advances and innovation in various fields of technology, and usually technologies whose development, practical applications, or both are still largely unrealised. Current examples of new and emerging technologies include artificial intelligence and robotics, raising new forms of impacts, both positive and negative, on human rights, and will likely only increase in their impacts in the future.

[7] UN General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35, 6 April 2018.

Module 2: Privacy and Free Expression in the Digital Environment

Module 2: Privacy and Free Expression in the Digital Age

Privacy in the digital age

As module 1 showed, the digital environment has opened up a number of opportunities for the promotion and protection of human rights, and acted as an enabling tool for human rights activists and defenders. However, it has also made individuals susceptible to surveillance on both an industrial and individualised scale, and created vulnerabilities that threaten their privacy and security.

In this module, we look at the standards for protecting the rights to privacy and freedom of expression. We then turn to key privacy- and free expression-related issues relevant to businesses in the digital environment, such as data protection, government requests for user information, and content moderation. Finally, we look at some of the best practices and tools that human rights defenders can use to engage with companies when it comes to the rights to privacy and free expression.

The right to privacy: an introduction

International law

While there is no universally accepted definition of "privacy" and therefore the right to privacy, the UN High Commissioner for Human Rights (UN HCHR) has said that it can be considered as "the presumption that individuals should have an area of autonomous development, interaction and liberty, a 'private sphere' with or without interaction with others, free from State intervention and from excessive unsolicited intervention by other uninvited individuals".[8]

A right to privacy has long been recognised in international law. Article 12 of the UDHR provides that "no one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks".

Similarly, Article 17 of the ICCPR provides that "no one shall be subjected to arbitrary or unlawful interference with his or her privacy, family, home or correspondence, nor to unlawful attacks on his or her honor and reputation," and that "everyone has the right to the protection of the law against such interference or attacks".

[8] United Nations High Commissioner for Human Rights, The right to privacy in the digital age, UN Doc. A/HRC/39/29, 3 August 2018, Para 5.

In its General Comment No. 16, the UN Human Rights Committee noted that "the obligations imposed by this article require the State to adopt legislative and other measures to give effect to the prohibition against such interferences and attacks as well as to the protection of this right".[9] It also states that an interference with the right to privacy is only permissible if it is neither arbitrary nor unlawful. In practice, this means that any interference must be provided for by law, be in pursuit of a legitimate aim, and be necessary and proportionate to achieving that aim.

More recently, in the context of the digital environment, the UN HCHR has noted that "informational privacy, covering information that exists or can be derived about a person and her or his life and the decisions based on that information, is of particular importance".[10] The combination of rapidly evolving digital technologies, the expansive applications of these technologies and the aggressive collection of personal information by states and businesses has magnified the points of tension over privacy, undermining its enforcement, and making it easier and faster for a wide array of actors to push boundaries.

Privacy from whom?

One of the first questions to ask when considering privacy in the digital environment, is from whom the privacy is being sought, whether at an individual or collective level. Before the development of the business and human rights framework, there was a simple answer: the state. As set out in module 1, under the framework, businesses also have a responsibility to respect the right to privacy. Sometimes, the two are connected: for example, when a government requests user information from a company. While these notes focus on the corporate responsibility to protect privacy, it is helpful to understand the difference between the roles and responsibilities of governments and companies in safeguarding privacy.

Privacy from government

The obligation on states to protect and respect the right to privacy is well-established. Article 2(1) of the ICCPR, for example, requires states to "respect and ensure" the rights in the ICCPR for all individuals within their territory and subject to their jurisdiction, without discrimination. However, the legal frameworks governing privacy-related issues (such as surveillance, or data protection) vary widely from country to country, and include different safeguards to ensure that government activities do not unduly infringe on individuals' privacy. While many

[9] UN Human Rights Committee, General Comment No. 16: Article 17 (Right to Privacy) The Right to Respect of Privacy, Family, Home and Correspondence, and Protection of Honour and Reputation, 1988.
[10] United Nations High Commissioner for Human Rights, "The right to privacy in the digital age", UN Doc. A/HRC/39/29, 3 August 2018, Para 5.

threats to privacy from the government do not involve the private sector at all, when it comes to the digital environment, the fact that much of the technology used by governments is developed by the private sector means that their role should also be considered. Some examples include the security of data and government databases, government surveillance, and the use of automated decisionmaking processes. Also, much of the private information sought by governments may be held by the private sector, for example the content of emails and messages transmitted on internet platforms.

Data security

There have been a number of major data breaches and scandals occurring in the tech sector over the past few years. Although a large amount of attention has been focused on how these breaches and hacks impact major tech companies and online platforms, there have also been a number of large data breaches involving governments.

One of the most globally discussed data breaches relates to the Aadhaar system, India's national identification system, which is the largest biometric database in the world. In 2015 and 2016, vulnerabilities were discovered that enabled anyone to download the private information of individuals in the database. In 2016 an investigation found that anyone could buy the information of over 1 billion Indians from this database for very little money.[11]

In the same year, the group Anonymous Philippines was able to hack into the website of the Philippines' Commission on Elections, deface it, and then call for stronger security measures on vote counting machines.[12] Shortly after, a second hacker group posted a link to what it claimed to be the entire database of the Commission on Elections. This was considered the biggest private data leak in the Philippines history, and it left millions of registered voters at risk.

Government surveillance

Many governments have instituted robust surveillance programmes (with some using them to target political opponents, dissenters, or minority groups) breaching or risking a breach of their citizens' right to privacy.

An example of a programme that poses risks to privacy is Section 702 of the Foreign Intelligence Surveillance Act in the US, which authorises the US government to target foreigners located abroad to collect foreign intelligence information. The law is controversial, since the definition of "foreign intelligence"

[11] BBC News, "Aadhaar: 'Leak' in world's biggest database worries Indians", 5 January 2018, available at: https://www.bbc.co.uk/news/world-asia-india-42575443.
[12] Hern, A., "Philippine electoral records breached in 'largest ever' government hack", *The Guardian,* 11 April 2016, available at: https://www.theguardian.com/technology/2016/apr/11/philippine-electoral-records-breached-government-hack.

is broad, and because it permits the warrantless collection of substantial quantities of US citizens' communications whenever they communicate with foreign targets.

Another example is the Social Credit System being developed by the Chinese government. By 2020, the system aims to provide each Chinese citizen with a credit score that is based on data collected by both private and public entities regarding their economic and social reputation. This is also intended to be applied to businesses. So far, citizens with low social credit scores have been banned from booking air and train travel and placed on travel blacklists. It is also predicted that the score will be used to bar children from certain schools. A positive social score on the other hand can provide greater access to loans and jobs.

Governments around the world are also increasingly adopting surveillance technologies such as facial recognition technology. The government of Zimbabwe, for example, recently formed a strategic partnership with Chinese artificial intelligence startup CloudWalk Technologies to import facial recognition software into the country that will initially be applied in law enforcement and security settings. The technology will eventually be applied in broader public use scenarios. [13] This has been particularly concerning given Zimbabwe's troubling human rights record in the past.

With the Aadhaar program in India, there have also been concerns that by pressuring citizens to sign up for an Aadhaar card and link all of their personal information, including telephone service, internet service, bank accounts, and so on to the system, it is creating a mass surveillance program that enables the government to monitor and keep track of every citizen. In addition, collecting large volumes of personal information increases the risks of data breaches and poses risks that databases will be targeted by hackers. Government databases like Aadhaar and the US' Office of Personnel Management database have been targeted by both private and foreign state attacks in the past.

Automated decision making

Another way in which government activities can impact privacy is through big data analysis using artificial intelligence. This allows states to identify patterns in the detailed information they collect about the lives of people, make inferences about their physical and mental characteristics, and create detailed personality profiles. Many of the systems used by governments are designed with the purpose of maximising the amount of information about individuals to analyse, profile, evaluate, categorise, and eventually make decisions, often automated, about them.

[13] Gwagwa, A., "Exporting Repression? China's Artificial Intelligence Push into Africa", *Council on Foreign Relations,* 17 December 2018, available at: https://www.cfr.org/blog/exporting-repression-chinas-artificial-intelligence-push-africa.

These activities can generate risks for individuals and societies. For example, big data analysis requires large datasets, and recent years have seen far-reaching data breaches exposing the people involved to identity theft and the disclosure of intimate information. In addition, the "scoring" and "grading" systems for individuals can be used to assess eligibility for medical care, other insurance coverage, financial services, and more. As we will discuss later, AI systems can also perpetuate bias and make it difficult for individuals to challenge decisions that affect them.

Privacy from companies

While the risks to privacy that can stem from government actions are significant, the focus of this guide is on the role of businesses. Ensuring that businesses respect the right to privacy requires efforts from both governments and the companies themselves. As noted in module 1, the obligation on states under international human rights law to ensure the protection of human rights also obliges them to ensure that the right to privacy is protected against the actions of businesses. This obligation is elaborated upon in Pillar 1 of the UNGPs, and states are expected to adopt a mix of measures, mandatory and voluntary, to ensure that businesses respect human rights, including the right to privacy. However, Pillar 2 of the UNGPs also sets out the responsibilities of the companies themselves to prevent and address adverse human rights impacts, including on the right to privacy. This responsibility exists independently of whether the state meets its own human rights obligations.

The specific actions that a company should take to respect the right to privacy will vary from company to company, however there are a range of initiatives that have developed recommendations to operationalise the UNGPs for companies in the digital environment. These include the Global Network Initiative's Principles on Freedom of Expression and Privacy (the GNI Principles)[14] and the Telecommunications Industry Dialogue Guiding Principles.[15] The Ranking Digital Rights Corporate Accountability Index evaluates a number of internet, mobile and telecommunications companies specifically on their disclosed commitments and policies affecting freedom of expression and privacy.[16]

Recent examples demonstrate how companies can adversely impact the right to privacy, such as data breaches of huge scope, exposing the persons concerned to identity theft and the disclosure of intimate information. These include the hack of LinkedIn in 2012, which was originally thought to have impacted 6.5 million

[14] Global Network Initiative, Principles on Freedom of Expression and Privacy of the Global Network Initiative, available at: https://globalnetworkinitiative.org/gni-principles.
[15] Telecommunications Industry Dialogue Guiding Principles, available at: http://www.telecomindustrydialogue.org/about/guiding-principles/.
[16] Ranking Digital Rights Corporate Accountability Index, available at: https://rankingdigitalrights.org/.

usernames.[17] In 2016, however, the group that orchestrated the hack revealed they had acquired the email and passwords of over 117 million LinkedIn users and were going to sell them.[18] Also in 2016, personal information belonging to approximately 57 million Uber customers and drivers, including drivers' licence numbers, was stolen by hackers.[19] The company kept the breach hidden for a year, and the company paid $100,000 to the hackers to cover up the breach.

These events seriously jeopardised the personal information and digital security of these platforms' users. And in many of these cases, the companies failed to disclose the breaches to the users, therefore leaving them, and their sensitive personal data, vulnerable to further manipulation and misuse. Under the UNGPs, companies have a responsibility to protect human rights, and the decision to not disclose the breaches demonstrates a failure to safeguard the right to privacy.

In addition, even without data breaches that leave user data vulnerable, companies can also intrude on the privacy of users by collecting vast troves of data and analyzing the data for patterns that may reveal highly personal information. This has been seen, for example, with the emergence of targeted ads. One striking case was when US department store Target assigned every customer a Guest ID number, tied to their credit card, name, or email address, and stored a history of everything that person bought together with any demographic information Target collected or bought from other sources. Using this data, Target was able to identify when women were pregnant, based on their buying habits, and send them targeted ads for items. This was often before the women or their families knew they were pregnant.[20]

Another example is smartphone makers, like Google, who have come under fire for their pervasive location tracking features that enable them to see everywhere a user has gone, often even when their phone is switched off.[21]

Privacy and data protection

Data protection is a particularly critical issue, and is only increasing in importance as the amount of data created and stored grows. There are a variety of safeguards

---

[17] BBC News, "LinkedIn passwords leaked by hackers", 7 June 2012, available at: https://www.bbc.co.uk/news/technology-18338956.
[18] BBC News, "Millions of hacked LinkedIn IDs advertised 'for sale'", 18 May 2016, available at: https://www.bbc.co.uk/news/technology-36320322.
[19] Wong, J. C., "Uber concealed massive hack that exposed data of 57m users and drivers", *The Guardian*, 22 November 2017, available at: https://www.theguardian.com/technology/2017/nov/21/uber-data-hack-cyber-attack.
[20] Hill, K., "How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did", *Forbes*, 16 February 2012, available at: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#439c2cad6668.
[21] Nakashima, R., "Google Tracks Your movements, Like it or Not," *Associated Press,* 16 August 2018, available at: https://apnews.com/828aefab64d4411bac257a07c1af0ecb.

that governments and businesses can implement to reduce the risks of privacy intrusions. One framework for protecting privacy is the Fair Information Practice Principles (FIPPs), which provides a set of best practices used by many different actors as guidance on implementing processes for information security and privacy. The preliminary version of the Principles was developed by the US Department of Health, Education, and Welfare in 1973, which the Organisation for Economic Co-operation and Development (OECD) built on to create the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, issued in 1980.[22]

The first edition of the OECD Guidelines was issued before many of the current privacy challenges that we face today, but versions of the FIPPs have been adopted by many governments and international organisations as part of their general frameworks for privacy protection. The US Federal Trade Commission has recommended their use by companies, as well as the government of Canada, the Council of Europe, the European Union, and others.
There are various formulations of the FIPPs, but the OECD formulation, which has been widely adopted, comprises the following principles:

- Collection limitation
- Data quality
- Purpose specification
- Use limitation
- Security safeguards
- Openness
- Individual participation
- Accountability

Privacy can be impacted by data collection, retention, security, and use, and each has its own particular considerations.

Data collection

Any company or organisation (including governments) that collects information about its users should follow clear data collection best practices. These ares necessary to protect individuals' privacy, but also to minimise what data could be lost in a data breach. Transparency around data collection can increase user trust in a product or service, and provide the necessary education for them to use the tools wisely.

Many organisations collect far more information than they actually need. Data is valuable to organisations for purposes of research, advertising, and

[22] OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, available at: https://www.oecd.org/internet/ieconomy/privacy-guidelines.htm.

implementation of new features, but much of the information collected is not necessary for the product or service to function. For example, when downloading certain applications to a smartphone, the app provider may ask for access to various types of personal information like a contacts list or camera roll images. They may also ask for access to sensors like the phone's microphone and camera. However, many of these applications will not require this supplementary information in order to function. To protect users' privacy, companies should practice data minimisation and limit the information they collect. Although a company may find access to your contacts list valuable so they can encourage you to engage with other contacts using the same app, that information isn't necessary and can pose significant privacy and security risks.

Organisations should allow users to control the data that is being collected about them, and disclose this information to users. In the example mentioned above, the application could ask the user whether it could access certain personal information on that smart phone. It is a data collection best practice to allow users to opt in or out of providing access to that information, rather than collecting it without user notification or permission.

Using tools like lights, sounds, or pop-ups to inform users when or how their information is being collected is a good best practice for organisations. This is especially relevant for physical devices, where features like cameras often have a light that turns on when they begin collecting video, notifying users that the camera has been engaged. Because there are so many instances in which cameras, audio recorders, or other sensors are harnessed to collect information on their users, clear notifications as to when those features are in use are important tools to protect user security and privacy as well as provide transparency around data collection practices.

Data retention

Encouraging companies to adhere to data retention best practices is a way to protect users from misuse or breach of their personal information. It is important that they disclose to users the timeline for deletion of different types of data like messages, photos, recordings, or accounts. Knowing these timelines can help users decide what type of information they want to use with the product or service, and allow them to better understand the risks posed by retention of that data. The more user data retained, the higher risks posed by data collection and storage, such as data breaches and other unauthorised access.

Companies should also only retain minimal user information, reducing the privacy and security risks that come with large-scale data collection and storage such as data breaches and other unauthorised access. The effects of a data breach such as those affecting the credit rating company Equifax or the US Government agency

Office of Personnel Management can be very harmful, and the more data companies have on their users, the higher the likelihood that they will be a target and that a breach would have catastrophic consequences.

People now tend to have many different types of devices, accounts, and applications, that are not always consistently used or maintained. It is best practice that, after a limited period of time and after warning inactive users, organisations close inactive accounts and delete user information. This puts that data at less of a risk for breach, and helps preserve security and privacy of users by giving fewer actors access to "abandoned" information they are no longer using.

Some countries or international agreements have guidelines or legal requirements surrounding data retention. For example, the European Union General Data Protection Regulation (GDPR) states that personal data should be kept for no longer than is necessary for the purposes for which it is being processed. There are some circumstances where personal data may be stored for longer periods (e.g. archiving purposes in the public interest, scientific,or historical research purposes).

Data security

Data security practices are extremely important for all products and services that collect information about their consumers. With the amount of personal data that many companies gain through the use of their products, they have an obligation to take steps to protect that data from accidental leaks or data breaches. There are a variety of best practice tools that can be used to protect users' security and privacy, as well as to reduce financial and reputational risk to the company or organisation.

- Encrypting data in transit and at rest can prevent breaches and information theft.
- Having a program to encourage vulnerability disclosure and manage hardware and software vulnerabilities allows organisations to benefit from the expertise of independent security researchers to help them reduce vulnerabilities in their product.
- Secure authentication practices, such as not using default passwords, requiring passphrases of a certain length and complexity, enabling multi factor authentication, and notifying users when account security settings have changed can help reduce the chances of a breach.
- Testing products or systems against known exploits helps reduce the chances of a serious vulnerability in the system.
- Regular patching, ideally with automatic updates, ensures that users benefit from the most secure version of the software.
- Having a transparent process for notifying users whose information has been breached increases user trust in the organisation.

- Implementing mitigation processes in case of a data breach helps increase user trust and prevent further security breaches.

Data use

Once a product or service has collected information about its users, ranging from biographical data like names and dates of birth, to biometric data like heartbeat and blood glucose levels, there are certain best practices it should follow regarding how it uses that data.

Whether it is sending someone "Happy Birthday" messages, optimising features to their needs, or sending them targeted information, companies or organisations sometimes use personal data in ways people do not consent to, or do not understand that they have already consented to. For example, personal data about a person's clicks on certain pop-ups, communications with friends, or age and gender, can be used to sell a user things specific to their assumed preferences. Companies make money by selling their users' data to advertisers who then benefit from that data to target them in order to better sell products. Organisations should disclose how they might use the information they collect from users so that they can make more informed choices about how they use that product or service. It is also a best practice for organisations to give users the ability to opt in or out of certain types of data use, like targeted advertising, and not use data in ways that the users have not consented to. This is good for protecting user security and privacy, as well as maintaining user trust and confidence in a service.

As many states debate their own privacy regime, a vibrant discussion is occurring about whether the above best practices are sufficiently protective. The GDPR's notice and consent model only goes so far, and there may be options for increasing those protections by either adapting the current regime or adopting new practices altogether. For example, adopting best practices concerning data portability, rectification, and data access, would expand user rights beyond those covered by the GDPR and could serve as a model for other countries to adopt. In addition, companies should incorporate safeguards for human rights because data collection practices and algorithmic decision making can lead to discriminatory treatment.

In January 2019 France's data privacy authority hit Google with a €50 million ($57 million USD) fine, claiming that the US search giant did not adequately explain to users how it handled their personal data, and did not properly obtain their permission for personalised ads.[23] The penalty is the largest so far under the GDPR. While the fine is relatively small for Google, legal analysts say it could mark the beginning of a raft of regulatory actions that will define how GDPR is interpreted.

[23] Fox, C., "Google hit with £44m GDPR fine over ads", *BBC News,* 21 January 2019, available at: https://www.bbc.co.uk/news/technology-46944696.

Under the law, EU regulators can fine companies up to 4 percent of their global annual revenue, or €20 million, whichever is larger.

Privacy and government requests for user data

Intelligence and law enforcement agencies have long sought data directly from companies, and, as internet and telecommunications companies collect and store more data, the number of requests has increased. As a result, it has now become common investigatory practice for these entities to request data from the services that a target may have used. If a company or organisation collects and stores user data, the likelihood of receiving such a legal request is high. As a result, it is vital that companies prepare for such requests and that users hold companies accountable when managing their data.

There are a range of domestic and international instruments that governments can use to justify and make requests for user data. For each one, the issuing authority and standard vary depending on the jurisdiction and what kind of information they are seeking. One common example is a Mutual Legal Assistance Treaty (MLAT), an agreement between two or more states that enables governments to seek information that is held by companies in another country for use in connection with their investigation and prosecution of criminal cases. For example, various governments may want to seek GMail information that is held by Google in the US. However, with the growth of the internet and expansion of electronic communications, MLATs have become cumbersome and time-consuming processes, and as a result many countries have sought to develop new legal frameworks to bypass the MLAT process. In 2018, the US enacted the CLOUD Act, which enables the US to enter into bilateral agreements with other countries that meet certain human rights standards, and once such an agreement is in place, the countries can send demands for electronic communications data directly to tech companies in the other country. The EU is considering an E-Evidence proposal that would set up a similar regime within the EU.

Dealing with government requests

There are a number of best practices companies can implement and users can push for in order to ensure the process of engaging with and responding to legal requests for user data is transparent and respectful of human rights.

The first best practice is to develop clear policies for processing and responding to government requests. This is a critical but challenging process and **must** involve consultations with legal counsel. Only a lawyer can help a company identify lawful requests and develop effective compliance mechanisms. There are a few broad considerations to ensure procedures and policies are transparent and efficient and that companies are pushing back on requests that are not appropriate.

- **Tracking requests and their status**: Companies should use a single, centralised process for tracking, tagging and keep tabs on the status of requests from the moment they are received until the time a response is provided to the government.
- **Reviewing and classifying requests for accuracy and validity**: Before a company can respond to a request, it has to identify the type of process and the agency or court that issued it. Companies should therefore have trained staff that can properly classify requests. This process is also vital for identifying requests that do not comply with legal requirements or would violate users' rights. and requests that companies should push back on. This is something the GNI Principles touch on when they highlight that "participating companies will respect and work to protect the privacy rights of users when confronted with government demands, laws or regulations that compromise privacy in a manner inconsistent with internationally recognised laws and standards..
- **Responding to requests**: Companies should work with their legal counsel to preemptively develop a formal process for responding to government requests. This will prevent errors and promote greater safeguarding of privacy and security.
- **Providing user notice**: Notifying users when their information has been requested is an important aspect of safeguarding user rights. However, at times the provision of notice can be delayed for ongoing investigations subject to applicable laws. Government requests are also increasingly accompanied by a nondisclosure order that prevents companies from informing the target of the investigation. For those requests without such orders, companies must decide whether, how, and under what circumstances they'll provide notice to their users. For requests with nondisclosure orders, companies must decide whether to challenge the order and/or to inform users after the order has been lifted.
- **Keeping data secure**: Information about law enforcement and intelligence requests is sensitive information in itself. In order to keep this information secure, companies should carefully consider how this data is maintained and who has access to it.
- **Challenging requests where necessary and appropriate**: Companies should have procedures in place that enable them to evaluate the validity and accuracy of requests and challenge requests that are out of scope, overbroad, or that infringe on privacy and security.

A public version of these policies should be posted online for users and other parties to view. This public edition of the policies should include information on what kind of legal orders and mechanisms the company accepts and responds to, the format in which requests must be made, the scope requests must fall under, and the user notification process. Providing a FAQ section that augments

understanding of how companies respond to government requests for user data is also helpful.

The second best practice for responding to legal requests for user data is for companies to publish regular and consistent transparency reports that provide quantitative and qualitative information on the scope and volume of government requests for user information they have received.

Transparency reporting is an extremely valuable practice that helps the public hold companies accountable to safeguard user data and enables companies to communicate with their users and lawmakers about privacy and security. In addition, transparency reporting provides a number of added benefits such as signaling company values, educating lawmakers, and easing customers' privacy concerns.

Although the amount of data the government is seeking has risen markedly, these requests still impact a very low percent of users. This is why in the wake of the Snowden disclosures, which suggested that major tech companies were turning over significant amounts of user data to the government tech companies, pressed governments to let them provide more information about the actual extent of government demands.

In 2016, New America's Open Technology Institute and Harvard University's Berkman Klein Center for Internet & Society published a Transparency Reporting Toolkit. The Toolkit surveyed how domestic internet and telecommunications companies were reporting on government requests for user data, and offered a set of guiding best practices on how their reporting can be improved going forward. These best practices focused on making these transparency reports clearer, more detailed, and more standardised across companies. Some of the best practices outlined in the Toolkit include: [24]

- **Reporting on different legal processes**: Companies should provide clear and granular categorizations of applicable legal processes. In addition, an ideal report will, at minimum, provide the number of government requests for each of these categories. In the US these categories include search warrants, subpoenas, other court orders, wiretap orders, pen register orders, and emergency requests.

- **Explaining legal processes**: Companies should provide clear and comprehensive explanations of legal processes. Definitions or a glossary explaining legal processes and other key terms used in the report can also

[24] Transparency Reporting Toolkit, available at: https://www.newamerica.org/oti/policy-papers/transparency-reporting-toolkit-reporting-guide-and-template/.

inform readers about the types of processes that might allow governments to access their data, while also helping to generate an understanding of the logistics behind transparency report, including how companies count legal processes.

- **Reporting on the subjects of requests and how users are impacted**: Companies should report the number of selectors (e.g., name, phone number, email address) specified in a request, including all unique identifiers . They should also report the number of users and/or accounts responsive to a request. Whether a company reports users vs. accounts depends on the services they offer. Companies should be clear in their reports about whether they are reporting on users, accounts, or both.

- **Reporting on the legal processes required for user information**: Because readers of transparency reports may be unfamiliar with the intricacies of applicable legal systems, it is important for companies to provide informative explanations of the legal processes the company requires in order to turn over specific types of user information.

- **Explaining "Content" and "Non-Content"**: While laws such as the US Electronic Communications Privacy Act define "content" and "non-content", it is important for companies to elaborate on these definitions so readers can understand the significance of the information in the transparency reports.

- **Reporting on outcomes and compliance with requests**: Companies should provide granular reporting on their compliance with requests for each kind of process (e.g., warrant, subpoena) and for each of the different ways a company may respond to a request (e.g., rejected, disclosed content).

- **Reporting on user notification**: Companies should provide clear, comprehensive, and granular reporting on notification of users specified in legal process requests. This includes reporting on three types of notifications: 1) When a request was under seal and the user could be notified, 2) When a request was not under seal and the user was notified, and 3) When a request was not under seal and the user was not notified.

These best practices apply to US-based companies and the relevant American legal procedures they engage with when managing government requests for user information. However, these best practices can be adapted to fit different countries and legal contexts.

Tools for advocates: encouraging best practices

Having looked at how the actions of companies and governments have a powerful impact on our privacy, it is important to look at what tools are available for human rights defenders to push governments to protect, and companies to respect, human rights. There are a number of best practice frameworks and models that can be used when attempting to assess and safeguard privacy online.

Getting Internet Companies to do the Right Thing

New America's Open Technology Institute's "Getting Internet Companies to do the Right Thing" report recognises the trends seen over the past decade (and highlighted in the previous module) in how companies have improved their privacy and security practices and what has led them to make these changes.[25] The report looks at case studies on companies in the US that are using transit encryption by default, offering two-factor authentication, and issuing transparency reports. The case studies outline the most common ingredients in a recipe for widespread adoption of a new privacy or security practice. For example, an initial crisis, like a major hack or data breach, that highlights the need for a best practice is a powerful tool that advocates can use to prompt a change in policy. These best practices can also be applied to the adoption of new free expression safeguards. In addition, although most of these milestones are applicable to US-based companies and an American policy context, they can be adapted for regionally-specific or country-specific cases.

The Digital Standard

The Digital Standard, a collective effort led by Consumer Reports, Disconnect, Ranking Digital Rights, and The Cyber Independent Testing Lab, with assistance from Aspiration, is a set of individual tests that taken together form a tool for evaluating the privacy and security impacts of a given piece of software or hardware.[26] It was created to define and reflect important consumer values that must be addressed in the development of software and hardware products. The Standard is underpinned by a set of guiding principles: internet connected devices and software-based products should be secure, consumer information should be kept private, ownership rights of consumers should be maintained, and products should be designed to combat harassment and help protect freedom of expression.

The Standard is composed of 35 different "tests" that can be used to measure products to see how their design and policies meet best practices for digital privacy and security. It also provides a model that companies can use to design and improve their products, ensuring that they are best in class on these issues and giving them an opportunity for product promotion in a crowded field. Companies

[25] New America, Getting Internet Companies to Do the Right Thing, available at: https://www.newamerica.org/in-depth/getting-internet-companies-do-right-thing/
[26] The Digital Standard, available at: https://www.thedigitalstandard.org/

are evaluated by a set of indicators under each test and must fulfill certain requirements to be certified as employing best digital security and privacy practices.

Who Has Your Back?

The Electronic Frontier Foundation has been releasing "Who Has Your Back?" reports since 2011.[27] The reports use a set of questions to evaluate US companies on how well they protect their customers and users, and include questions on government requests for user information, policies on data retention, and transparency reporting. Other organisations have used some version of EFF's Who Has Your Back? (government requests version) metrics to evaluate the practices of companies in their own countries, adapting the categories to local laws.

What's next? Privacy and emerging technologies

Emerging technologies pose a whole host of new privacy and security challenges.

Artificial Intelligence

The intelligence community is using artificial intelligence (AI) to help manage the exponential increase of data it are collecting. Computer analytics are used to read and understand the data in order to free up humans for more specific tasks. Using AI, intelligence agencies are auditing routine functions that analysts, curators, and collectors have manually done in the past. For example, the US National Geospatial-Intelligence Agency is using AI to sort large quantities of image data, helping the Agency to pull out specific targets such as enemy safehouses and airfields. As in other types of data analysis, this frees up humans who used to do this type of work manually to perform more complex tasks that cannot be performed by a machine.

However, with the rising use of AI there is also a potential for unintended consequences with issues of privacy, transparency, safety, control, and bias. For example, artificial intelligence can exhibit dataset bias, association bias, interaction bias, automation bias, and confirmation bias. These can enter the system as the result of simple mistakes or lack of oversight in data aggregation techniques, and also due to the nature of machine learning algorithms, which will perpetuate unknown biases in the data used to train the systems. There are movements to make ethical guides for AI, and it is a best practice for governments and other organisations to stress test their AI models and test for bias. Some of these best practices include making sure there are human analysts included in any

[27] For the most recent report, see Electronic Frontier Foundation, Who Has Your Back? 2019, available at: https://www.eff.org/wp/who-has-your-back-2019

AI data analysis to watch for potential bias, and to balance the need for innovation with the obligation to benefit and safeguard society.

Internet of Things

There are three major concerns that arise out of expanded data collection through the Internet of Things (IoT) that can be addressed by the implementation of best practices for privacy and security. First, these massive datasets provide opportunities for exploitation by nefarious actors, and create risks of general privacy violations through poor privacy and security best practices. When information about users' very private lives is being collected on such a massive scale it creates a significant target for hackers and cybercriminals. This information is also often transmitted over insecure connections, potentially exposing private information. There was a 600% increase in attacks against IoT devices from 2016-2017, and implementing security measures like encryption, strong authentication practices, and testing software against known vulnerabilities are best practices that can help protect these massive troves of user data against hackers.

Second, there is the potential that companies collecting data may misuse it for the companies' own benefit, in ways that are not beneficial or comprehensible to the users of those devices. Companies can use data for targeted advertising, or they could sell data to other companies for profit. Consumers don't necessarily understand the implications of giving manufacturers and developers all of this private user data, and so best practices like clear and comprehensive Terms of Service and privacy policies can help protect consumers and allow them to make informed choices.

Third, the IoT devices themselves can be harnessed to commit harmful acts. When surveillance cameras, routers, internet doorbells, thermostats, and other infrastructure sensors are IoT-enabled, there is potential for these devices to be harnessed in DDOS attacks or even exploited to make them malfunction in ways that could hurt their users. Security best practices like encryption, authentication, and patching are best practices that can help protect users.

Freedom of expression in the digital age

Over recent years, many countries have enacted laws and regulations imposing liability on online platforms for the content that they host (or "online content regulation"). Although the aims of these legal rules are generally to try to control hate speech, disinformation, and abuse, the rules can ultimately interfere with individuals' rights to freedom of expression. Separately from government regulation, making decisions about online content (or "online content moderation") has, for the most part, been undertaken by platforms, who are

increasingly responding to this challenge through the use of outsourced workers or artificial intelligence. These companies face competing pressures to ensure their platforms are safe while also respecting the freedom of expression of their users.

At present, online content moderation is based on national legislation (each company is required to comply with the local law wherever it does business) as well as on platforms' Community Guidelines or Terms of Service, a set of private rules that outline what is allowed and what is not allowed on their platforms.

The right to freedom of expression: an introduction

International law

Freedom of expression can broadly be defined as an individual's right to say, write, and produce content on almost anything they'd like without restriction. On an international level, freedom of expression is enshrined in Article 19 of the UDHR as well as the ICCPR. Article 19(2), read together with Article 2 of the ICCPR, provides for state parties' obligations to respect and ensure "the right to freedom of expression," which includes the "freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice".

In its General Comment No. 34, the UN Human Rights Committee highlighted that they are "indispensable conditions for the full development of the person", "essential for any society" and that "they constitute the foundation stone for every free and democratic society". It also outlines that according to Article 19 of the ICCPR, states should ensure that restrictions on expression, online or otherwise, are lawful, necessary, and proportionate.

Paragraph 3 of Article 19 expressly states that the exercise of the right to freedom of expression carries with it special duties and responsibilities. For this reason there are only a limited number of reasons for taking actions that may impinge on free expression rights, namely actions taken to respect the rights or reputations of others, or those taken to protect national security or public order (*ordre public*) or public health or morals.

These international treaties and agreements generally outline the obligations for governments regarding human rights, whereas frameworks such as the UNGPs provide guidance for how companies should integrate respect for human rights.

Freedom of expression: protection for what and from whom?

Governments and companies engage with freedom of expression online in different ways. In the US, for example, governments generally cannot restrict speech due to the First Amendment, but since companies are private entities, they

are more free to impose limits. This is why companies are able to engage in content regulation practices and establish their own Terms of Service and content standards.

Challenges to free expression are growing around the world, both from companies and governments. Some examples of controversies around free expression include:

- The passage of the Network Enforcement Act (or NetzDG law) in Germany in June 2017. The law mandates that social media companies of a particular size must remove "obviously illegal speech" such as terror content that is flagged to them within 24 hours or face massive fines.
- The August 2018 reporting alleging that Google was attempting to launch a censored search engine under the name Project Dragonfly in China.
- The spread of disinformation on Facebook in Myanmar over the past few years that stoked ethnic and religious tensions and conflict.
- The spread of fake news stories via WhatsApp in India that has resulted in mob violence and the deaths of dozens of people since 2017.
- A string of cybersecurity and fake news bills in countries like Malaysia, Kenya, and Egypt.

In order to safeguard free expression, governments have an obligation to ensure that any laws they pass that restrict speech comply with international human rights norms.

Although companies are not similarly bound by the international human rights framework, they nonetheless have a responsibility under the UNGPs to protect the right to freedom of expression. Therefore, companies should ensure their speech standards are compatible with the international human rights framework and are clearly communicated to users. Furthermore, they should conduct due diligence on requests received through government or private processes to take down content, in order to ensure these processes and requests are just, fair, and do not result in overbroad censorship.

Freedom of expression and governments

States have a duty to promote and protect the free exercise of the right to freedom of opinion and expression. States may restrict the right to freedom of expression under very limited circumstances. Under Article 19(3) of the ICCPR, restrictions on the right to freedom of expression must be "provided by law", and necessary "for respect of the rights or reputations of others" or "for the protection of national security or of public order (*ordre public*), or of public health and morals". State

obligations with respect to restrictions on online expression are the same as those offline.[28]

Article 19(3) imposes a three-part test for permissible restrictions on freedom of expression:
First, restrictions must be "provided by law". The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has noted in the past his concern that restrictions on freedom of expression should be subject to regular legislative process, including the participation of the interested persons through public comment processes and public hearings).[29] In evaluating the standard provided by law, the UN Human Rights Committee has noted that any restriction "must be made accessible to the public" and "formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly".[30] Moreover, it "must not confer unfettered discretion for the restriction of freedom of expression on those charged with its execution".[31]

Second, restrictions must only be imposed to protect legitimate aims, which are limited to those specified under Article 19(3), that is "for respect of the rights or reputations of others" or "for the protection of national security or of public order (*ordre public*), or of public health and morals". The term "rights (...) of others" under Article 19(3)(a) includes "human rights as recognized in the Covenant and more generally in international human rights law".[32]

Third, restrictions must be necessary to protect one or more of those legitimate aims. The requirement of necessity mandates an assessment of the proportionality of restrictions, with the aim of ensuring that restrictions "target a specific objective and do not unduly intrude upon the rights of targeted persons".[33] The restriction must be "the least intrusive instrument among those which might achieve the desired result".[34]

---

[28] UN General Assembly, Resolution 68/167, The right to privacy in the digital age, UN Doc. A/RES/68/167, 21 January 2014; UN Human Rights Committee, General comment No. 34: Article 19: Freedoms of opinion and expression, UN Doc. CCPR/C/GC/34, 12 September 2011.
[29] UN Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, UN Doc. A/HRC/29/32, 22 May 2015.
[30] UN Human Rights Committee, General comment No. 34: Article 19: Freedoms of opinion and expression, UN Doc. CCPR/C/GC/34, 12 September 2011.
[31] UN Human Rights Committee, General comment No. 34: Article 19: Freedoms of opinion and expression, UN Doc. CCPR/C/GC/34, 12 September 2011.
[32] UN Human Rights Committee, General comment No. 34: Article 19: Freedoms of opinion and expression, UN Doc. CCPR/C/GC/34, 12 September 2011.
[33] UN General Assembly, Report of the Special Rapporteur on the promotion and the protection of the right to freedom of opinion and expression, David Kaye, UN Doc. A/70/361, 8 September 2015.
[34] UN Human Rights Committee, General comment No. 34: Article 19: Freedoms of opinion and expression, UN Doc. CCPR/C/GC/34, 12 September 2011.

The Human Rights Council has also advised that states should "refrain from imposing restrictions on: discussions on government policies and political debate; reporting on human rights, government activities and corruption; engaging in election campaigns, peaceful demonstrations or political activities; and expression of opinion and dissent, religion or belief, including by persons belonging to minorities or vulnerable group".[35]

The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has explained that to comply with the criteria of Article 19(3), "States should only seek to restrict content pursuant to an order by an independent and impartial judicial authority, and in accordance with due process and standards of legality, necessity and legitimacy".[36] States should also "refrain from imposing disproportionate sanctions, whether heavy fines or imprisonment, on Internet intermediaries, given their significant chilling effect on freedom of expression".[37]

The UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression has also urged states to "refrain from adopting models of [online content] regulation where government agencies, rather than judicial authorities, become the arbiters of lawful expression. They should avoid delegating responsibility to companies as adjudicators of content, which empowers corporate judgment over human rights values to the detriment of users".[38] Instead, "[s]mart regulation, not heavy-handed viewpoint-based regulation, should be the norm, focused on ensuring company transparency and remediation to enable the public to make choices about how and whether to engage in online forums".[39] In the US, the First Amendment limits the extent to which the government can enact regulations setting rules for what content is permitted.

Freedom of expression and companies

As use of technology platforms such as Facebook, Twitter, and Google have expanded, these companies have gained more influence and power, and have assumed the role of digital public squares, as well as de facto gatekeepers of online speech.

---

[35] UN Human Rights Council, Resolution 12/16, Freedom of opinion and expression, UN Doc. A/HRC/RES/12/16, 12 October 2009.
[36] UN General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35, 6 April 2018.
[37] UN General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35, 6 April 2018.
[38] UN General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35, 6 April 2018.
[39] UN General Assembly, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35, 6 April 2018.

However, as increasing numbers of jurisdictions have enacted laws regulating online content, companies have had to adjust their policies and procedures to comply with each of these local laws. For example, in Germany it is illegal to deny the Holocaust and as a result platforms that host user-generated content must treat such speech as illegal in the context of that country. As a result Facebook hired 10,000 extra content moderators to ensure compliance with the legislation and has been actively enforcing the regulation on its platforms.

In addition to ensuring compliance with national legislation, companies also develop their own standards and regulations around permissible speech on their platforms. These rules may restrict certain categories of speech beyond what would or could be prohibited by law in particular countries. For example, Facebook prohibits nudity, violence, bullying, and harassment, even though some speech in these categories would be protected by the First Amendment if the US government sought to prohibit it.

Companies both have a commitment to comply with speech-related laws around the world, and as private actors, have the right to develop their own standards for permissible content on their services within the bounds of these laws.

Grappling with these competing pressures whilst managing their individual impact on free expression is an ongoing challenge for many companies. As they increasingly face pressure to remove content—including hate speech and terror content—in order to make their platforms safer, tensions between the public, the company, and regulators regularly come to a head and have resulted in open disagreement.

Although the removal of all harmful content may be a tempting response to regulatory and public pressure, overzealous censorship of speech is a significant human rights concern and threat to free expression. Companies should carefully assess these competing concerns, and companies need to be held accountable in their management of online speech.

Freedom of expression and content regulation/moderation

What is content moderation?

Content moderation can be defined as the process companies employ to review and potentially remove, restrict, or regulate content that has been identified or flagged as violating a particular law or a company's content policies or Terms of Service. Content moderation can also result in the suspension or closure of the accounts of users who violate these regulations.

Just as companies use their privacy policies to outline how data is managed and governed on their platforms, they use Terms of Service to define standards and norms of acceptability on their platforms, including for speech. In this context, Terms of Service are also commonly referred to as Community Guidelines or Community Standards.

Case study: Facebook's Community Standards

In May 2018, Facebook released a detailed edition of their Community Standards, which outline standards of acceptability for content on the platform. The Standards also include a policy rationale and a clear list of what not to post.

Here is an example of what they share for their policy on promoting or publicising crime:

"Policy Rationale: We prohibit people from promoting or publicising violent crime, theft, and/or fraud because we do not want to condone this activity and because there is a risk of copycat behavior. We also do not allow people to depict criminal activity or admit to crimes they or their associates have committed. We do, however, allow people to debate or advocate for the legality of criminal activities, as well as address them in a rhetorical or satirical way."

They also outline that you should not post content that depicts, admits, or promotes the following criminal acts committed by you or your associates:

- Acts of physical harm committed against people
- Acts of physical harm committed against animals except in cases of hunting, fishing, religious sacrifice, or food preparation/processing
- Poaching or selling endangered species or their parts
- Staged animal vs. animal fights
- Theft
- Vandalism or property damage
- Fraud
- Trafficking
- Sexual violence or sexual exploitation, including sexual assault

What kinds of content are removed?

Some of the categories of content that are most commonly targeted for removal online are:

Content that is illegal in a particular region or country: Content that is illegal in one region or country but not others is often not removed from the platform

altogether; rather, it is restricted or geo-blocked so that it cannot be viewed by users in the country or region where it is illegal. Some examples of categories of content that are dealt with in this manner are instances of Holocaust denial in Germany where it is illegal, insults directed at the royal family in Thailand, and speech that "insults Turkishness" in Turkey. Typically content of this kind is removed as a result of legal requests by governments or others who flag the content as illegal. Companies can also proactively identify such content using automated tools.

Content that infringes on intellectual property rights: Companies often receive requests from users, other companies,and at times, governments to remove content that infringes on intellectual property rights. This includes copyright and trademark infringing content, and at times patent and counterfeit related claims, although the latter two are not as frequent. Most American companies process copyright claims as per the Digital Millennium Copyright Act, unless there is a local copyright law in a particular country that is applicable. The same goes for trademark infringement.

Copyright and trademark requests are generally thought of as the most straightforward types of requests, but nuance and context is needed for any form of request that is submitted to companies, and it's important to be cognisant of how these policies can impact the right to free expression.

For example, there have been some cases where intellectual property related takedown request procedures have been misused by governments in an attempt to quell dissent and silence activists and opposition members. This happened in Ecuador in 2015, when the government allegedly hired a Spanish law firm to submit DMCA takedown notices to companies like Google, Twitter, and Vimeo on behalf of state officials targeting critical documentaries, tweets, and search results that included images of those officials, alleging copyright infringement. Because removal of this category of content was generally thought of as straightforward, many of these requests were processed. However, some of the removed content was restored after the posters filed counter-notices. This is just an example of how these procedures can be abused, and therefore if you are a content creator online, and find that you're having your free expression challenged in such a manner, it is vital that you are aware of the appeal mechanisms that are available to you.

Content covered by the "right to be forgotten": In May 2014, the Court of Justice of the European Union ruled that citizens of the EU could ask that search engines delist search results tied to their names if the information in the result was "inadequate, irrelevant or excessive in relation to the purpose of processing".[40]

[40] *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González,* C-131/12, 13 May 2014.

This right to erasure was also included in the EU's General Data Protection Regulation which came into effect in May 2018. A similar law came into effect in Russia in January 2016. At present, the law is only applied to search engines, and as a result only search engines such as Google and Microsoft's Bing receive these requests. However, these takedowns subsequently impact content on other platforms that appear in the search engine results.

Content prohibited under Community Guidelines: The Community Guidelines of a platform define standards of acceptability for content and often touch on issues including graphic violence, nudity, terror content, hate speech, bullying, harassment, etc. Companies like Facebook have taken steps to release detailed versions of their guidelines, in order to promote greater transparency around their practices and their thinking on policy implementations. Still, there is a lot of room to grow and expand this reporting, especially across the industry. There is still a large amount of ambiguity around how companies moderate content based on their guidelines, and what their guidelines are.

At what point is content removed?

There are common points at which content can be removed in its lifecycle: before content is published, after it has been flagged by a user, and as a result of an active internal assessment of content being posted on platforms (whether the assessment is conducted by algorithms or human moderators).

Transparency reporting

As we've seen, companies remove content across a number of content categories. This significantly impacts free expression and free speech online. In 2018, New America's Open Technology Institute released a Transparency Reporting Toolkit that surveyed how 35 global and domestic internet and telecommunications companies were reporting on six categories of content takedowns in their transparency reports.[41] The report also offers a set of guiding best practices on how their reporting can be improved going forward, with a focus on making them clearer, more detailed, and more standardized across companies.

The toolkit showed there is a general tendency to report on government requests and copyright requests, but there is less reporting for the other forms of content removals. OTI also found that when a major company takes the first step, it is usually to push other competitors to adopt similar practices. For example, Google released a comprehensive transparency report on its Terms of Service based takedowns for YouTube and shortly after Facebook followed suit.

[41] Transparency Reporting Toolkit, available at: https://www.newamerica.org/oti/policy-papers/transparency-reporting-toolkit-reporting-guide-and-template/.

The last couple of years have seen a critical mass of companies issue transparency reports on their content takedowns, share their content policies, and expand mechanisms like appeals processes for takedowns. Some of the best practices highlighted in the Toolkit are:

- **Issuing regular reports following clearly and consistently delineated reporting periods**: Companies should issue transparency reports on a consistent timeline, and clearly and consistently delineate the reporting period for each issued report. Currently, there is no industry-wide standard for how often companies should publish reports. However, because reports issued more often and covering shorter periods can offer more granular information, the best practice is to publish quarterly, if practical.

- **Issuing reports specific to the type of demand**: By reporting separately on different types of demands or takedowns (e.g., government and other legal demands, copyright requests, trademark requests etc.), rather than lumping them all together, a company is able to describe the breadth of demands they have received and the volume and impact of each of these takedown categories.

- **Reporting on types of demands using specific numbers**: By reporting statistics that separately describe the number of demands a company receives over a given time period for each different type of takedown, companies can show which types of demands are most common. The best practice is to report on the types of demands using specific numbers; percentages alone are not sufficient (though they are a helpful supplement), nor are numeric ranges.

- **Breaking down demands by country**: In order to demonstrate the geographic scope of demands a company is receiving, and to show which countries' governments or laws are most actively restricting online free expression, companies should specify how many requests originate from each specific country. The most effective way to provide this information is to create a list or map of all countries relevant to a company's operations and indicate the number of demands received from each.

- **Reporting on categories of objectionable content targeted by demands**: By reporting on the categories of objectionable content targeted by different types of content demands, a company can outline the varying reasons that parties are asking for content to come down, and also indirectly demonstrate the relative prevalence of different types of problematic content on their services.

- **Reporting on products targeted by demands**: Some companies maintain and

support multiple products. By specifying which of a company's products are being targeted by which demands, a transparency report can better reflect how those demands are impacting the range of its offerings, highlight differences in impact between its services, and better enable comparisons of that impact with other companies' comparable services.

- **Reporting on specific government agencies/parties that submitted demands**: By reporting on which specific government agencies or entities submitted content-related demands, a company can describe which elements of government in which countries are the most active in seeking to police online content, which can in turn help identify misuse or overuse of authority or actions outside of a particular part of a government's jurisdiction, as well as overall trends in what content which parts of government are targeting.

- **Specifying which laws pertain to specific demands**: Because most major internet and telecommunications companies operate in multiple countries, it is important for companies to report on which laws and legal frameworks govern user speech and communications and lead to takedown requests.

- **Reporting on the number of accounts and items specified in demands**: Companies should report on the number of accounts *and* items specified in demands as this enables a better understanding of the full breadth of those demands.

- **Reporting on the number of accounts and items impacted by demands**: Companies should also report on the number of accounts *and* items impacted by demands. Such reporting offers the most direct measure of how many speakers and how much free expression is being silenced as a result of demands (and how many/how much is being effectively defended by the company). It also enables a comparison of the requested impact versus the actual impact, which in turn offers a greater understanding of both the quality and legality of the requests being made and the company's rates of compliance with those requests.

- **Reporting on how the company responded to demands**: Reporting on how a company responds to requests across different issue areas is vital for understanding how companies comply with legal frameworks, government demands and user requests. In addition, it also illustrates the role companies play in protecting or censoring speech.

Tools for advocates: encouraging best practices

Having looked at how the actions of companies and governments have a powerful impact upon our freedom of expression, it is important to look at what tools are available for human rights defenders to push governments to protect and companies to respect human rights. There are a number of best practice frameworks and models that can be used when attempting to assess and safeguard freedom of expression online.

The Santa Clara Principles on Transparency and Accountability in Content Moderation

During the first Content Moderation and Removal at Scale conference in Santa Clara in 2018, several organisations, advocates, and academic experts who support the right to free expression online convened a small private workshop to consider how best to obtain meaningful transparency and accountability around internet platforms' increasing moderation of user-generated content.

At the second Content Moderation at Scale conference these organizations and individuals released the Santa Clara Principles, which consist of three principles. The principles outline initial steps that companies engaged in content moderation should take to provide meaningful due process to impacted speakers and better ensure that the enforcement of their content guidelines is fair, unbiased, proportional, and respectful of users' rights. These principles cover three key aspects: numbers, notice, and appeals.

These principles were meant to serve as a starting point, outlining minimum levels of transparency and accountability to serve as the basis for a more in-depth dialogue in the future.

The recommendations outlined in the Santa Clara Principles are:

Numbers
Companies should publish the numbers of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines. At a minimum, this information should be broken down along each of these dimensions:

- Total number of discrete posts and accounts flagged.
- Total number of discrete posts removed and accounts suspended.
- Number of discrete posts and accounts flagged, and number of discrete posts removed and accounts suspended, by category of rule violated.
- Number of discrete posts and accounts flagged, and number of discrete posts removed and accounts suspended, by format of content at issue (e.g., text, audio, image, video, live stream).

- Number of discrete posts and accounts flagged, and number of discrete posts removed and accounts suspended, by source of flag (e.g., governments, trusted flaggers, users, different types of automated detection).
- Number of discrete posts and accounts flagged, and number of discrete posts removed and accounts suspended, by locations of flaggers and impacted users (where apparent).
- This information should be provided in a regular report, ideally quarterly, in an openly licensed, machine-readable format.

Notice
Companies should provide notice to each user whose content is taken down or account is suspended about the reason for the removal or suspension. In general, companies should provide detailed guidance to the community about what content is prohibited, including examples of permissible and impermissible content and the guidelines used by reviewers. Companies should also provide an explanation of how automated detection is used across each category of content. When providing a user with notice about why a post has been removed or an account has been suspended, a minimum level of detail includes:

- URL, content excerpt, and/or other information sufficient to allow identification of the content removed.
- The specific clause of the guidelines that the content was found to violate.
- How the content was detected and removed (flagged by other users, governments, trusted flaggers, automated detection, or external legal or other complaint). The identity of individual flaggers should generally not be revealed, however, content flagged by governments should be identified as such, unless prohibited by law.
- Explanation of the process through which the user can appeal the decision.
- Notices should be available in a durable form that is accessible even if a user's account is suspended or terminated. Users who flag content should also be presented with a log of content they have reported and the outcomes of moderation processes.

Appeal
Companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension. Minimum standards for a meaningful appeal include:

- Human review by a person or panel of persons that was not involved in the initial decision.

- An opportunity to present additional information that will be considered in the review.
- Notification of the results of the review, and a statement of the reasoning sufficient to allow the user to understand the decision.
- In the long term, independent external review processes may also be an important component for users to be able to seek redress.

What's next? Freedom of expression and emerging technologies

As emerging technologies develop they are increasingly being applied to the management of online content. AI mediated content moderation is the best, and most common, example of this. As we previously discussed, many platforms deploy AI systems in order to identify and regulate content at scale. This occurs at various stages of moderation including pre-publication moderation and during active assessment of content on the platform.

Artificial Intelligence

AI systems vary in their ability to identify problematic content. This depends both on the AI system and on the type of content, with some types of content being far harder to identify than others. For example, some models are very accurate when identifying child sexual abuse materials on the platform while others struggle to identify instances of hate speech.

AI tools have resulted in some contentious takedowns. For example, the famous photograph of "Napalm Girl" depicting a little girl running away from a napalm strike during the Vietnam War was removed by Facebook's content moderation algorithms for violating the platform's nudity policy—without any appreciation for historical value and weight.

Other unintended victims of AI-mediated content moderation have been human rights groups who repost images of terrorist atrocities. Various such groups have had their content taken down for violating platform's policies on terror content and propaganda, despite their intent being the opposite, namely to raise awareness rather than recruiting. This happened for example in Myanmar, where Rohingya activists had their content removed and accounts shut down.

As a result of the problems with AI-mediated content moderation, companies must consistently involve human moderators in all content decision making practices. Although humans can also make errors, human moderators can add cultural context and understanding and help prevent such mistakes and such instances of curtailed free expression, from taking place.

Recently, civil society and advocates have pushed companies for greater transparency around how their automated tools are being deployed in content moderation. In Google's most recent transparency report on the enforcement of

their Community Standards, for example, they show that the majority of content they removed was flagged by automated tools. However, they don't provide breakdowns for what categories of content, like terror content, hate speech etc. these automated flags targeted.

One positive step we've seen with regard to providing greater transparency around these efforts is that Facebook has begun disclosing in its Community Standards Enforcement Report how much of the content it actioned was appealed by users, and how much actioned content the platform restored as a result of appeals and other reasons.

Module 3: The UN Guiding Principles Pillar I: the State Duty to Protect Human Rights

Module 3: UN Guiding Principles Pillar I: the State Duty to Protect Human Rights

While much of the focus of the business and human rights framework is on the responsibility of businesses to respect human rights, the entire first pillar of the UNGPs looks at the duty of *states* to ensure that human rights are respected by businesses. There has been much concern, however, over the lack of action taken by states so far. The UN Working Group on the issue of human rights and transnational corporations and other business enterprises (the Working Group on Business and Human Rights) has said that governments "are not fulfilling their duty to protect human rights, either failing to pass legislation that meets international human rights and labour standards, passing legislation that is inconsistent or failing to enforce legislation that would protect workers and affected communities".[42]

What's the reason for this failure? One likely factor is that governments either do not know, or do not fully understand, the range of options available to them in developing a "smart mix of measures" (to use the wording of the Commentary to the UNGPs) which would help them fulfil their duty. This module looks at the requirements of the first pillar of the UNGPs, with a focus on its application to companies in the digital environment, and sets out some of the steps that governments can take to protect human rights from business impacts. That first pillar of the UNGPs sets out a series of foundational and operational principles relating to the state duty to protect human rights.

Foundational principles

Principle 1: States must protect against human rights abuse within their territory and/or jurisdiction by third parties, including business enterprises. This requires taking appropriate steps to prevent, investigate, punish and redress such abuse through effective policies, legislation, regulations and adjudication.

As a first step, the Working Group on Business and Human Rights and the International Corporate Accountability Roundtable have recommended the assessment of existing laws as an initial step and several governments have taken this step.

[42] UN General Assembly, Report of the Working Group on the issue of human rights and transnational corporations and other business enterprises, UN Doc. A/73/163, 16 July 2018.

Principle 2: States should set out clearly the expectation that all business enterprises domiciled in their territory and/or jurisdiction respect human rights throughout their operations.

The most effective way to do this is through clear, available and enforced legislation. Guidance, engagement and encouragement are also critical elements of the process. As such, many National Action Plans on Business and Human Rights (NAPs) - which are looked at in this module later - have focused on this element to date.

Operational principles

General regulatory and policy functions of the state

Principle 3: In meeting their duty to protect, States should:

(a) Enforce laws that are aimed at, or have the effect of, requiring business enterprises to respect human rights, and periodically to assess the adequacy of such laws and address any gaps;
(b) Ensure that other laws and policies governing the creation and ongoing operation of business enterprises, such as corporate law, do not constrain but enable business respect for human rights;
(c) Provide effective guidance to business enterprises on how to respect human rights throughout their operations;
(d) Encourage, and where appropriate require, business enterprises to communicate how they address their human rights impacts.

The types of laws that could require business enterprises to respect human rights include, for example, consumer protection, environmental protection, criminal and civil liability. To ensure that laws and policies do not constrain but enable business respect for human rights, states need to ensure the cost of being a business, and doing business, doesn't incentivise rights abuses. They must also provide effective guidance to business enterprises on how to respect human rights throughout their operations. Beyond law, this is about proactive assistance. And, to encourage, and where appropriate require, business enterprises to communicate how they address their human rights impacts. Here, transparency is key.

The state-business nexus

Principle 4: States should take additional steps to protect against human rights abuses by business enterprises that are owned or controlled by the State, or that receive substantial support and services from State agencies such as export credit agencies and official investment insurance or guarantee agencies, including, where appropriate, by requiring human rights due diligence.

States should take additional steps to protect against human rights abuses by business enterprises that are owned or controlled by the state, or that receive substantial support and services from state agencies such as export credit agencies and official investment insurance or guarantee agencies, including, where appropriate, by requiring human rights due diligence.

The UNGPs are clear that states have special responsibilities for the conduct of state-owned enterprises. This principle also addresses state-assistance to private businesses such as export credits, investments, insurance, etc. For example, Canada has committed to tie the provision of these services to demonstrations of respect for human rights.

Principle 5: States should exercise adequate oversight in order to meet their international human rights obligations when they contract with, or legislate for, business enterprises to provide services that may impact upon the enjoyment of human rights.

This principle primarily has to do with concessions, e.g., mineral extraction, oil and gas operations, hydroelectric facilities, transportation, utilities. The ability to exercise adequate oversight in this regard can also be a condition for external support for projects by international financial institutions like the World Bank.

Principle 6: States should promote respect for human rights by business enterprises with which they conduct commercial transactions.

This principle includes ways to use procurement as an incentive for promoting respect for human rights. For example the US State Department requires companies bidding on large, private security contracts to be certified against human rights standards and members of a multistakeholder, human rights initiative, the "International Code of Conduct on Private Security Service Providers Association".

Supporting business' respect for human rights in conflict-affected areas

Principle 7: Because the risk of gross human rights abuses is heightened in conflict-affected areas, states should help ensure that business enterprises operating in those contexts are not involved with such abuses, including by:

(a) Engaging at the earliest stage possible with business enterprises to help them identify, prevent and mitigate the human rights-related risks of their activities and business relationships;
(b) Providing adequate assistance to business enterprises to assess and address the heightened risks of abuses, paying special attention to both gender-based and sexual violence;
(c) Denying access to public support and services for a business enterprise that is involved with gross human rights abuses and refuses to cooperate in addressing the situation;
(d) Ensuring that their current policies, legislation, regulations and enforcement measures are effective in addressing the risk of business involvement in gross human rights abuses.

Ensuring policy coherence

Principle 8: States should ensure that governmental departments, agencies and other State-based institutions that shape business practices are aware of and observe the State's human rights obligations when fulfilling their respective mandates, including by providing them with relevant information, training and support.

Principle 9: States should maintain adequate domestic policy space to meet their human rights obligations when pursuing business-related policy objectives with other States or business enterprises, for instance through investment treaties or contracts.
Principle 10: States, when acting as members of multilateral institutions that deal with business-related issues, should:

(a) Seek to ensure that those institutions neither restrain the ability of their member States to meet their duty to protect nor hinder business enterprises from respecting human rights;
(b) Encourage those institutions, within their respective mandates and capacities, to promote business respect for human rights and, where requested, to help States meet their duty to protect against human rights abuse by business enterprises, including through technical assistance, capacity-building and awareness-raising;
(c) Draw on these Guiding Principles to promote shared understanding and advance international cooperation in the management of business and human rights challenges.

National Action Plans on Business and Human Rights

National Action Plans (NAPs) are state policy strategies setting out an overarching strategy and concrete activities to address a specific policy issue or issues. NAPs have long been used by governments to define and address a wide array of issues, from human trafficking to health, to climate change, to women, peace, and security.

In the business and human rights context, a National Action Plan on Business and Human Rights has been defined as "an evolving policy strategy developed by a State to protect against adverse human rights impacts by business enterprises in conformity with the UN Guiding Principles on Business and Human Rights".[43]

The UN Working Group's Guidance on NAPs sets out a series of criteria for NAPs:[44]

- A NAP should be founded on the UNGPs, meaning that is has to be based on international human rights standards and reflect the UNGPs' emphasis on state obligations and business responsibilities. In order to ensure this, the guidance recommendations that states conduct capacity building on the UNGPs within government, and takes the UNGPs as a guiding instrument (with underlying international instruments) when identifying and deciding on measures to address protection gaps.
- NAPs should respond to specific challenges in the national context. In other words, one size will not fit all. States should identify and map adverse human rights impacts occurring in the country's territory as well as abroad by companies domiciled in the country, conduct and update an assessment of state and business implementation of the UNGPs including the implementation of existing law, regulations and voluntary initiatives, and focus on addressing concrete impacts when drafting the document.
- NAPs should be developed transparently and inclusively. This means that they should involve as many relevant government entities as possible, consult and take into account the views and needs of non-governmental stakeholders throughout the process of NAP development, monitoring and update, outline and update a clear time plan on the NAP process, and share information and results of assessments and consultations with all relevant stakeholders on a regular basis.
- NAPs should be regularly reviewed and updated. The state should commit to an open-ended process in the early stages, clarify in the NAP when an existing NAP will be updated, and provide clear timelines for the implementation of actions defined in NAPs and measure progress.

[43] UN Working Group's guidance on NAPs, available at: https://www.ohchr.org/Documents/Issues/Business/UNWG_NAPGuidance.pdf.
[44] UN Working Group's guidance on NAPs, available at: https://www.ohchr.org/Documents/Issues/Business/UNWG_NAPGuidance.pdf.

In terms of process, the guidance sets out a five phase process for creating, implementing, and updating NAPs. These are:

- Initiation;
- Assessment and consultation;
- Drafting;
- Implementation; and
- Update.

Note that there are different views about some elements of the process. For example, the Danish Institute and the International Corporate Accountability Roundtable (a coalition of corporate accountability focused NGOs) stresses the need for national baseline assessments for NAPs, but not all governments have employed this practice. Only 6 of 21 NAPs published as of late 2018 conducted baseline assessments. In addition, the US government based its 2017 NAP on both the UNGPs and the OECD Guidelines for Responsible Business Conduct.

Initiation

Steps that the Working Group flags as key to initiating a NAP include:

- Seeking formal commitment from the government to a NAP
- Creating a format for cross-departmental collaboration and designate leadership within government
- Create a format for engagement with non-governmental stakeholders
- Develop and publish a work plan and allocate adequate resources

Examples: National human rights institutions and academic institutions have raised awareness, conducted research and laid the groundwork for NAPs in Philippines, Ghana, Korea, Malaysia, Morocco, and South Africa. In Switzerland and the Netherlands, the parliaments called on the governments to develop NAP and Spain has a work plan which was early published and updated frequently.

Assessment and consultation

The state should identify and catalogue the main adverse human rights impacts created by businesses as well as any gaps in government and corporate responses. "Adverse impacts" include any impacts occurring on the state's territory as well as abroad (with the involvement of a company domiciled in the country). In this step it is considered key to consult stakeholders on priorities and concrete actions to include in the NAP. The guidance suggests two criteria: 1) severity of human rights impacts; and 2) leverage of government to make change on the ground.

Drafting

At this stage in the process, the government plays a key role in ensuring participation, mediating interests, and ensuring coherence. The guidance provides a roadmap for this stage by providing an outline, a set of underlying principles for the NAP and a non-exhaustive list of measures to consider when consulting on each guiding principle.

Many governments have chosen to conduct consultations on the draft prior to finalising and launching it, using this as a key opportunity to raise awareness of business and human rights issues in the country, and to engage business on their responsibilities.

Examples of good practice:

- Finland, Spain, Switzerland invited written feedback from civil society organisations and companies
- India has published a zero draft for consultation
- Colombia launched a public consultations with stakeholders on a draft

A central element of NAPs is defining a government response to adverse human rights impacts created by business. The guidance recommends incorporating four principles when drafting:

1. **Focus on addressing concrete impacts when drafting the document**: The selection of the impacts to be addressed with priority should follow two key criteria: 1) the severity of adverse human rights impacts and 2) the leverage of the government in bringing about change on the ground;
2. **Use the UNGPs to identify how to address impacts**: Governments should rely on the UNGPs to identify specific and achievable measures on how to prevent, mitigate and redress adverse human rights impacts by business enterprises. At the same time, governments should refer to the UNGPs addressing businesses in pillars II and III. In particular, the concept of human rights due diligence should be promoted as the thread ensuring coherence in the government's activities outlined in NAPs.
3. **Identify a 'smart mix' of mandatory and voluntary, international and national measures.**
4. **Ensure effective protection from gender specific impacts**: This includes integrating a gender analysis to identify such impacts, including by collecting gender disaggregated data, and committing to measures which prevent, mitigate and allow for the remediation of gender-based impacts.

Implementation

The implementation of the NAP will be facilitated if, for each action outlined in the NAP, clear objectives, responsibilities, and timelines are defined and if the necessary financial resources are made available. Multistakeholder monitoring groups may also be very useful to create independent monitoring frameworks (e.g. Finland's Committee on Corporate Social Responsibility). Finally, it is good practice to have a government focal point that can respond to requests and concerns from non-governmental stakeholders.

Update

It is good practice to:

- Include date for evaluation and update in the NAP
- Have and evaluation by an independent entity, such as NHRI, and consult with relevant stakeholders
- Update assessments of adverse impacts and protection gaps
- Inform and consult with stakeholders in preparation for updated NAP

It is important to remember that NAPs are meant to be iterative and evolving policy guidance tools, rather than a static set of commitments. Best practice is for NAPs to be renewed every two or three years.

Timeline of National Action Plans

The 2011 "Communication" by the European Commission inviting member states to conduct NAPs sparked a flurry of NAPs being issued by EU member states. Since then, other regional and multilateral bodies have encouraged and, in some instances, facilitated the development of NAPS, including the AU, ASEAN, the Council of Europe, and the G7 and G20, and the OAS.

Specific national approaches:

There are a number of some NAPs that make specific reference to ICT sector, such as the NAPs issued by the ones from Czech Republic, Finland, France, Lithuania, Luxembourg, the Netherlands, Poland, Spain, Sweden, Switzerland, United Kingdom, and the United States.

In December 2011, the Institute for Human Rights and Business and Shift were selected by the European Commission to develop sector-specific guidance on the corporate responsibility to respect human rights in the ICT sector. This process resulted in the publishing of the ICT sector guide on implementing the UN Guiding

Principles on Business and Human Rights. And in part, it explains the focus on European countries developing standards for ICT within their NAPs.

Sweden: The Swedish government highlights in its NAP that internet freedom and privacy are among the great global issues of the future. It states that it is fundamental for Sweden that the human rights that apply offline, also apply online. The NAP claims that as a result of a Swedish initiative, the OECD Guidelines for Multinational Enterprises now call on companies to support human rights on the internet.

Ireland: Similarly to Sweden, the Irish NAP highlights past actions such as providing a fourfold increase in the funding for the work of the Data Protection Commission in recent years; and the UK NAP highlights that it has strengthened international rules relating to digital surveillance, including leading work in the Wassenaar Arrangement to adopt new controls on specific technologies of concern.

The Netherlands: The Netherlands undertook a Sector Risk Analysis in 2014 which identified the electronics sector as among those with the greatest risk of adverse human rights impacts. The government has committed to negotiating voluntary corporate social responsibility agreements that focus on transparency, dialogue with stakeholders, and monitoring of agreements with those sectors.

Other countries have developed specific action points.

Poland: The Polish NAP commits the government to draft a regulation to counteract restrictions on the freedom of speech. In particular it developed rules governing the liability of internet intermediaries for hate speech and freedom of speech, stating that: "The Ministry of Digital Affairs plans to draft a regulation to counteract restrictions on the freedom of speech, on the one hand, and to block illegal content on the Internet, on the other. Legislative work is being carried out that clarifies the procedure for notice and takedown of the illegal content online, as well as strengthens legal safeguards for freedom of speech in the activities of electronic service providers. These efforts address i.a. issues related to hate speech or incitement to violence, as well as the use of unauthorised technical restrictions on freedom of speech in social media".

Finland: The Finnish NAP has proposed to create a roundtable discussion on how to ensure the protection of privacy in Finland with the government agencies, ICT companies and civil society, stating that: "The right to privacy, protection of personal data, and the protection of confidential messages are fundamental human rights." and proposes roundtable on privacy with government, companies, and civil society.

Switzerland: The government of Switzerland developed regulations of technologies for internet and mobile communication surveillance, noting that "Technologies for Internet and mobile communication surveillance can be used for both civilian and military purposes, i.e. they are dual-use goods. They can be an element in state repression, for example, thereby exposing the business enterprises that manufacture or trade in them to an increased risk of becoming involved in human rights abuses. The export or brokerage of technologies for Internet and mobile communication surveillance is governed by goods control legislation. (...) The transfer of intellectual property, including expertise and the grant of rights, concerning technologies for Internet and mobile communication surveillance was also made subject to license."

United Kingdom: The government of the United Kingdom has developed two NAPs. In the 2013 NAP they were committed to develop guidance on ICT exports with impact on human rights and freedom of expression online. In particular it states that "The Government will do the following to reinforce its implementation of its commitments under Pillar 1 of the UNGPs: (v) In line with the UK Cyber Exports Strategy, develop guidance to address the risks posed by exports of information and communications technology that are not subject to export control but which might have impacts on human rights including freedom of expression on line." That guide, "Assessing Cyber Security Export Risks: Human Rights and National Security" was published in November 2014.

The United Kingdom's 2016 Updated NAP also makes a reference to ICT in the section discussing Actions taken:

"To give effect to the UN Guiding Principles, the Government has: (...) strengthened international rules relating to digital surveillance, including leading work in the Wassenaar Arrangement to adopt new controls on specific technologies of concern. Specifically, new controls were agreed on: – equipment and software for creating and delivering "intrusion software" designed to be covertly installed on devices to extract data. – "Internet surveillance systems" which can monitor and analyse Internet traffic and extract information about individuals and their communications."

United States: The United States NAP commits to develop a regular mechanism to identify, document, and publicise lessons learned and best practices related to corporate actions that promote and protect human rights online. And to foster continued engagement among relevant stakeholders to support ongoing dialogue and collaboration on respecting human rights within the ICT sector. It does so by recognising that "The impact and importance of business conduct in the ICT sector has grown as social, commercial, educational, and recreational interactions increasingly take place online. State, working with other agencies and stakeholders, will develop a regular mechanism to identify, document, and

publicise lessons learned and best practices related to corporate actions that promote and protect human rights online. State will also foster continued engagement among relevant stakeholders to support ongoing dialogue and collaboration on respecting human rights within the ICT sector".

Attempts to establish a state duty to protect against human rights abuses by business

In addition to NAPs, there have been various other legally binding and non-binding efforts to mandate a state duty to protect against human rights abuses by business. These include domestic laws and regulations that require corporate due diligence or human rights reporting:

- **Argentina**: The Balancing Social and Environmental Responsibility Act in Buenos Aires requires companies that have had their main business in Buenos Aires for over one year and have over 300 employees to prepare an annual report of their social, environmental and economic impact.
- **Nigeria**: The Petroleum Industry Bill requires all licensed petroleum operators in Nigeria to submit an environmental quality management plan that complies with environmental laws and establishes and monitors health and safety standards in the industry. It also sets out obligations with respect to human rights, including labour rights and gender equality.
- **India**: Section 135 of the Indian Companies Act requires businesses with revenues over 10bn rupees to spend 2% of their profits on corporate social responsibility initiatives.
- **France:** The Duty of Vigilance Law requires companies to develop and implement a "Vigilance Plan" every year, and to report on the prior year's plan.

There are also a series of other non-binding mechanisms:

- **OECD Guidelines for Multinational Enterprises**: The Guidelines are an annex to the OECD Declaration on International Investment and Multinational Enterprises. They were adopted in 1976 and subsequently revised. Human rights were included in the 2011 update. Recommendations addressed by governments to multinational enterprises operating in or from adhering countries. They provide non-binding principles and standards for responsible business conduct in a global context consistent with applicable laws and internationally recognised standards.
- **OECD National Contact Points**: National Contact Points (NCPs) for Responsible Business Conduct (RBC) promote the Guidelines, respond to enquiries and provide a mediation and conciliation platform to help resolve cases of alleged non-observance of the Guidelines (known as "specific

instances"). 48 governments have committed to create an NCP for RBC, 35 OECD countries and 13 non-OECD countries. Over 400 cases have been handled by NCPs.

- **OECD Due Diligence Guidance**: This guidance, published in May 2018, provides practical support to enterprises on the implementation of the OECD Guidelines for Multinational Enterprises. There is also due diligence guidance by sector (e.g., Minerals, Extractive, Garment and Footwear, etc.); the ICT sector has no specific due diligence guidance.
- **Open Government Partnership (OGP)**: Open Government Partnership brings together government reformers and civil society leaders to create action plans that make governments more inclusive, responsive and accountable. Participating governments agree to create an OGP national action plan. OGP has "Five Grand Challenges" around which countries may choose to develop an action plan. One of those is "Increasing Corporate Accountability," in which concrete commitments include measures that address corporate responsibility on issues such as the environment, anti-corruption, consumer protection, and community engagement.

Module 4: The UN Guiding Principles Pillar II: the Corporate Responsibility to Respect Human Rights

Module 4: UN Guiding Principles Pillar II: The Corporate Responsibility to Respect Human Rights

The second pillar of the UNGPs – Respect – requires businesses to refrain from creating "adverse human rights impacts" wherever, and however, they do business. In effect, this pillar introduces a positive obligation on companies to comply with the international human rights framework regardless of the legal requirements in the country they are doing business in. Even in countries where the government doesn't fully comply with their own human rights duties, companies are obliged to understand their human rights impacts under the international framework and take concrete and proactive steps to address any adverse ones.

Specifically, the commentary for Principle 11 explains that corporate respect "requires taking adequate measures for their prevention, mitigation and, where appropriate, remediation". In spite of this guidance, according to human rights benchmarking and rating assessments, the majority of companies are not meeting the requirements set by the UNGPs. This failure is particularly marked in the digital sector - with the latest Corporate Accountability Index by Ranking Digital Rights (RDR) scoring only a few companies above 50 percent on their commitment to human rights and governance structure.

Foundational principle

Principle 11: Business enterprises should respect human rights. This means that they should avoid infringing on the human rights of others and should address adverse human rights impacts with which they are involved.

Principle 12: The responsibility of business enterprises to respect human rights refers to internationally recognized human rights – understood, at a minimum, as those expressed in the International Bill of Human Rights and the principles concerning fundamental rights set out in the International Labour Organization's Declaration on Fundamental Principles and Rights at Work.

Principle 13: The responsibility to respect human rights requires that business enterprises:

(a) Avoid causing or contributing to adverse human rights impacts through their own activities, and address such impacts when they occur;

(b) Seek to prevent or mitigate adverse human rights impacts that are directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts.

Principle 14: The responsibility of business enterprises to respect human rights applies to all enterprises regardless of their size, sector, operational context, ownership and structure. Nevertheless, the scale and complexity of the means through which enterprises meet that responsibility may vary according to these factors and with the severity of the enterprise's adverse human rights impacts.

Principle 15: In order to meet their responsibility to respect human rights, business enterprises should have in place policies and processes appropriate to their size and circumstances, including:

(a) A policy commitment to meet their responsibility to respect human rights;
(b) A human rights due diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights;
(c) Processes to enable the remediation of any adverse human rights impacts they cause or to which they contribute.

There are a number of initiatives through which companies have operationalised these principles:

- **The Voluntary Principles on Security and Human Rights**: Established in 2000, the Voluntary Principles on Security and Human Rights are a set of principles designed to guide companies in the extractive sector in maintaining the safety and security of their operations within an operating framework that encourages respect for human rights. Participants in the Voluntary Principles Initiative — including governments, companies, and NGOs — agree to proactively implement or assist in the implementation of the Voluntary Principles.
- **International Code of Conduct Association**: The purpose of the Association is to promote, govern and oversee implementation of the International Code of Conduct and to promote the responsible provision of security services and respect for human rights and national and international law in accordance with the Code. The Code includes a wide range of standards and principles for the responsible provision of private security services which can be broadly summarised in two categories: first, principles regarding the conduct of Member Company personnel based on international human rights and humanitarian law standards including rules on the use of force, sexual violence, human trafficking and child labour; and second, principles

regarding the management and governance of Member Companies including the selection, vetting and proper training of personnel.

- **Know the Chain**: KnowTheChain is a resource for companies and investors to understand and address forced labor risks within their global supply chains. Through benchmarking current corporate practices and providing practical resources that enable companies to operate more transparently and responsibly, KnowTheChain drives corporate action while also informing investor decisions. KnowTheChain is committed to helping companies make an impact in their efforts to address forced labor.
- **Corporate Human Rights Benchmark:** Corporate Human Rights Benchmark Ltd, is a not for profit company created to publish and promote the Corporate Human Rights Benchmark. The Corporate Human Rights Benchmark was launched in 2013 as a multi-stakeholder initiative and draws on investor, business and human rights, and benchmarking expertise from 7 organisations.

There are also a series of Initiatives that specifically focus on business and human rights in the digital environment:

- **Global Network Initiative (GNI)**: The Global Network Initiative was launched in 2008 as a multistakeholder platform. GNI was the product of more than two years of deliberation by companies, human rights and press freedom organisations, academics, and investors. GNI participants work together in two mutually supporting ways. The GNI Principles ("the Principles") and Implementation Guidelines. GNI Global Principles on Freedom of Expression and Privacy: The Global Principles on Freedom of Expression and Privacy have been developed by companies, investors, civil society organisations and academics who aim to protect and advance freedom of expression and privacy in the Information and Communications Technology (ICT) industry globally. The Principles are based on internationally recognised laws and standards for human rights and their application is informed by the UN Guiding Principles on Business and Human Rights. GNI Implementation Guidelines: GNI's Implementation Guidelines provide details on how participating companies should put the GNI Global Principles into practice. The purpose of the document is to a) describe a set of actions by which a company would demonstrate that it is implementing the Principles with improvements over time, and b) provide companies with direction and guidance on how to implement the Principles.
- **Ranking Digital Rights Corporate Accountability Index**: Ranking Digital Rights produces a Corporate Accountability Index that evaluates the world's most powerful internet, mobile, and telecommunications companies on their disclosed commitments and policies affecting freedom of expression and privacy. The RDR Index provides a roadmap for how internet, mobile, and telecommunications companies—as well as other companies

throughout the sector—can improve how and what they disclose about policies and practices that affect digital rights. The RDR Index builds on the UN Guiding Principles on Business and Human Rights and on the Global Network Initiative principles and implementation guidelines. It draws on a body of emerging global standards and norms around data protection, security, and access to information and establishes benchmarks and ranks companies on disclosed policies and practices affecting their users' freedom of expression and privacy rights.

- **Who has your back?:** In this annual report, the Electronic Frontier Foundation examined the policies of major internet companies — including ISPs, email providers, cloud storage providers, location-based services, blogging platforms, and social networking sites — to assess whether they publicly commit to standing with users when the government seeks access to user data.

Operational principles

Policy commitment

Principle 16: As the basis for embedding their responsibility to respect human rights, business enterprises should express their commitment to meet this responsibility through a statement of policy that:

(a) Is approved at the most senior level of the business enterprise;
(b) Is informed by relevant internal and/or external expertise;
(c) Stipulates the enterprise's human rights expectations of personnel, business partners and other parties directly linked to its operations, products or services;
(d) Is publicly available and communicated internally and externally to all personnel, business partners and other relevant parties;
(e) Is reflected in operational policies and procedures necessary to embed it throughout the business enterprise.

When it comes to privacy freedom of expression, and tech companies responsibility to reflect their commitment to human rights through "operational policies and procedures", Terms of Service and privacy policies are critically important.

Terms of Service usually contain information about:

- Who may use the service;
- Overview of content and conduct policies;

- Legal text such as warranties, disclaimers, liability

When it comes to terms of service and privacy policies, these should (at a minimum) comply with three key standards. They should be: (a) clear; (b) easy to understand; and (c) accessible.

The Ranking Digital Rights Corporate Accountability Index methodology sets out further expectations from tech companies so as to demonstrate their compliance with Principle 16 of the UNGPs:

- **Corporate-level commitment to freedom of expression and privacy**: We expect companies to make an explicit statement affirming their commitment to freedom of expression and privacy as human rights, and to demonstrate how these commitments are institutionalised within the company. Companies should disclose clear evidence of: senior-level oversight over freedom of expression and privacy, and employee training and whistleblower programs addressing these issues; human rights due diligence and impact assessments to identify the impacts of the company's products, services, and business operations on freedom of expression and privacy; systematic and credible stakeholder engagement, ideally including membership in a multi-stakeholder organisation committed to human rights principles, including freedom of expression and privacy; a grievance and remedy mechanism enabling users to notify the company when their freedom of expression and privacy rights have been affected or violated in connection with the company's business, plus evidence that the company provides appropriate responses or remedies.
- **Terms of Service and privacy policies**: We expect companies to provide Terms of Service agreements and privacy policies that are easy to find and understand, available in the primary languages of the company's home market, and accessible to people who are not account holders or subscribers. We also expect companies to clearly disclose whether and how they directly notify users of changes to these policies.
- **Terms of Service enforcement:** We expect companies to clearly disclose what types of content and activities are prohibited, and their processes for enforcing these rules. We also expect companies to publish data about the volume and nature of content and accounts they have removed or restricted for violations to their terms, and to disclose if they notify users when they have removed content, restricted a user's account, or otherwise restricted access to content or a service.
- **Handling user information**: We expect companies to disclose what information they collect, what information they share, the types and names of the third parties with whom they share it, the purpose for collecting and sharing user information, and their data retention policies. Companies should also provide clear options for users to control what information is

collected and shared, including for the purposes of targeted advertising, and should clearly disclose if and how they track people across the web using cookies, widgets, or other tracking tools embedded on third-party websites. We also expect companies to clearly disclose whether users can obtain all public-facing and internal data they hold, including metadata.

- **Handling of government and private requests**: We expect companies to clearly disclose their process for responding to government and private requests to restrict content and user accounts and to hand over user information. We expect companies to produce data about the types of requests they receive and the number of these requests with which they comply. Companies should notify users when their information has been requested and disclose if laws or regulations prevent them from doing so.
- **Identity policies**: We expect companies to disclose whether they ask users to verify their identities using government-issued ID or other information tied to their offline identities. The ability to communicate anonymously is important for the exercise and defense of human rights around the world. Requiring users to provide a company with identifying information presents human rights risks to those who, for example, voice opinions that do not align with a government's views or who engage in activism that a government does not permit.

Human rights due diligence

Principle 17: In order to identify, prevent, mitigate and account for how they address their adverse human rights impacts, business enterprises should carry out human rights due diligence. The process should include assessing actual and potential human rights impacts, integrating and acting upon the findings, tracking responses, and communicating how impacts are addressed. Human rights due diligence:

(a) Should cover adverse human rights impacts that the business enterprise may cause or contribute to through its own activities, or which may be directly linked to its operations, products or services by its business relationships;
(b) Will vary in complexity with the size of the business enterprise, the risk of severe human rights impacts, and the nature and context of its operations;
(c) Should be ongoing, recognizing that the human rights risks may change over time as the business enterprise's operations and operating context evolve.

Principle 18: In order to gauge human rights risks, business enterprises should identify and assess any actual or potential adverse human rights impacts with

which they may be involved either through their own activities or as a result of their business relationships. This process should:

(a) Draw on internal and/or independent external human rights expertise;
(b) Involve meaningful consultation with potentially affected groups and other relevant stakeholders, as appropriate to the size of the business enterprise and the nature and context of the operation.

Principle 19: In order to prevent and mitigate adverse human rights impacts, business enterprises should integrate the findings from their impact assessments across relevant internal functions and processes, and take appropriate action.

(a) Effective integration requires that:
(i) Responsibility for addressing such impacts is assigned to the appropriate level and function within the business enterprise;
(ii) Internal decision making, budget allocations and oversight processes enable effective responses to such impacts.

(b) Appropriate action will vary according to:
(i) Whether the business enterprise causes or contributes to an adverse impact, or whether it is involved solely because the impact is directly linked to its operations, products or services by a business relationship;
(ii) The extent of its leverage in addressing the adverse impact.

Principle 20: In order to verify whether adverse human rights impacts are being addressed, business enterprises should track the effectiveness of their response. Tracking should:

(a) Be based on appropriate qualitative and quantitative indicators;
(b) Draw on feedback from both internal and external sources, including affected stakeholders.

Principle 21: In order to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders. Business enterprises whose operations or operating contexts pose risks of severe human rights impacts should report formally on how they address them. In all instances, communications should:

(a) Be of a form and frequency that reflect an enterprise's human rights impacts and that are accessible to its intended audiences;

(b) Provide information that is sufficient to evaluate the adequacy of an enterprise's response to the particular human rights impact involved;
(c) In turn not pose risks to affected stakeholders, personnel or to legitimate requirements of commercial confidentiality.

Due diligence, in its broad terms, is the standard of care a person or organisation will take before entering into an agreement with another party in order to assess risks. In business, this often applies before major investments or when pursuing a new business venture, such as a new product or expanding operations. Human rights due diligence assesses the human rights risk to people in its first instance and therefore a risk to the company as well. Human rights due diligence should happen continually to assess how risks change or develop over time.

Although due diligence and risk management systems are more complex in larger organisations, tech SMEs can often have large and diverse customer bases which spread over a wide geographical area. This can raise the risk to human rights and therefore increases the need for human rights due diligence.

Human Rights Impact Assessments (HRIAs) should include:

- A review of the UDHR, UNGPs, local laws, and the human rights environment;
- Risk scenarios based on the company's particular services or operations;
- Consultations with stakeholders.

There is no one-size fits all approach to HRIAs. Stakeholder engagement and remedy can present unique challenges for the ICT sector given the scope of the user base for many of these companies.

Resources on how to conduct HRIAs include:

- **Business and Human Rights Resource Center**: The Business and Human Rights Resource Center website includes a range of resources on HRIAs, including tools and guidance for conducting HRIAs, examples of implementation of community-led HRIAs, project- and group-led HRIAs, research and analysis on HRIAs, etc.
- **Conducting an Effective Human Rights Impact Assessment: Guidelines, Steps, and Examples, BSR**: Business for Social Responsibility published a report in March 2013 that provides helpful insight into how businesses can conduct human rights impact assessments. The report provides an overview of what constitutes a human rights impact assessment (HRIA) as well as guidelines and in-practice examples of HRIAs that BSR has assisted companies in conducting.

- **Human Rights Compliance Assessment, Danish Institute for Human Rights**: The Danish Institute for Human Rights offers an impact assessment tool that incorporates a database of 195 questions and 947 indicators, each measuring the implementation of human rights in company policies and procedures. The "Quick Check" tool, which is publicly available, focuses on human rights in relation to employment practices, community impact, and supply chain management.

Remediation

Principle 22: Where business enterprises identify that they have caused or contributed to adverse impacts, they should provide for or cooperate in their remediation through legitimate processes.

Principle 23: In all contexts, business enterprises should:

(a) Comply with all applicable laws and respect internationally recognized human rights, wherever they operate;
(b) Seek ways to honour the principles of internationally recognized human rights when faced with conflicting requirements;
(c) Treat the risk of causing or contributing to gross human rights abuses as a legal compliance issue wherever they operate.

Principle 24: Where it is necessary to prioritize actions to address actual and potential adverse human rights impacts, business enterprises should first seek to prevent and mitigate those that are most severe or where delayed response would make them irremediable.

Module 5: The UN Guiding Principles Pillar III: Access to Remedy

Module 5: UN Guiding Principles Pillar III: Access to Remedy

Article 8, UDHR

Everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law.

When it comes to access to remedy under the international human rights framework, Article 8 of the UDHR made clear what might already seem obviously, namely that it's not enough simply to prohibit human rights violations; where one does take place, something should be done to help remediate the victim. However in doing so, Article 8 makes clear that the document should not just be considered as a list of rules for states, but something practical and effective for individuals.

Article 2(3), ICCPR:

(a) Each State Party to the present Covenant undertakes (...) to ensure that any person whose rights or freedoms as herein recognised are violated shall have an effective remedy, notwithstanding that the violation has been committed by persons acting in an official capacity.

(b) To ensure that any person claiming such a remedy shall have his right thereto determined by competent judicial, administrative or legislative authorities, or by any other competent authority provided for by the legal system of the State, and to develop the possibilities of judicial remedy

(c) To ensure that the competent authorities shall enforce such remedies when granted.

When the ICCPR was developed, it also included provisions on access to remedy, and went further than the UDHR in a number of respects. First, it guarantees a remedy to the human rights contained within the ICCPR, not merely those which are protected by a state's constitution or national laws. Second, it makes clear that a person is entitled to an effective remedy, even if the violation was committed by a person "acting in an official capacity"; as such, the fact that a particular violation was permitted by national law, or carried out by a person following lawful orders, will not be a defence. Third, it provides that any person who considers that their human rights have been violated should be able to bring a claim to a competent

judicial, administrative or legislative authority. Fourth, that authority should be able to provide, and enforce, a remedy.

The UN Human Rights Committee has gone further in explaining what Article 2 requires. In its General Comment No. 31, the Committee made a number of points:

- First, the positive obligations on states to ensure the rights in the ICCPR can only be fully discharged if individuals are protected by the state, not just against violations of ICCPR rights by its agents, but also against acts committed by private persons or entities that would impair the enjoyment of ICCPR rights in so far as they are amenable to application between private persons or entities.
- Second, remedies must not only be effective, but accessible (rather than theoretical).
- Third, there is no one single form of remedy which will make it effective, and different remedies will be needed to address the violations of different groups.
- Fourth, the remedy must amount to "reparation" i.e. it must as closely as possible put the person in the situation they would have been in, had the violation not occurred. While this will usually involve compensation, it can take other forms as well.

Building on these general obligations under international human rights law, the UNGPs dedicate the third pillar solely to the issue of access to remedy, both by states and businesses.

Foundational principle

Principle 25: As part of their duty to protect against business-related human rights abuse, States must take appropriate steps to ensure, through judicial, administrative, legislative or other appropriate means, that when such abuses occur within their territory and/or jurisdiction those affected have access to effective remedy.

Principle 25 focuses on what states must do, and says that as part of their duty to protect against business-related human rights abuses, they must take certain steps to ensure that when human rights abuses occur, that those affected have access to an effective remedy.

There are two parts of this principle that are worth exploring a little further. First, it provides for a range of different ways that access to remedy can be ensured - including judicial and administrative procedures, as well as through legislation. Second, the human rights abuse has to occur in that state's territory or jurisdiction.

This can make things difficult when it comes to the tech sector since the harm might be suffered in one state, but the cause stems from a company or an individual based in another.

Interestingly, the commentary of Principle 25 says that its overall aim is "to counteract or make good any human rights harms that may have occurred" which sounds like simply putting the person back to the position they would have been in had the harm not occurred. But it then goes on to say that it might in fact mean more than this. This could include apologies, whether in public or private or guarantees of non-repetition.

Ensuring access to remedy also requires access to grievance mechanisms, and the UNGPs set out three types of grievance mechanisms that exist: first, and the most common, are state-based judicial mechanisms such as courts and tribunals. Second, there are state-based non-judicial mechanisms. There are many examples of these across the world, including ombudspersons or national human rights institutions. Third, there are also non-state-based grievance mechanisms, such as a company's own complaints procedures, but also regional courts. In all cases, the procedures for the provision of remedy should be impartial, protected from corruption and free from political or other attempts to influence the outcome. Effective grievance mechanisms should not only be easily accessible to the public but they should also be fully explained to ensure complete understanding of the process and affected rights holders should be actively encouraged to use such mechanisms. This will create a sense of trust that their grievances will be dealt with.

There is no exhaustive list of what kinds of remedies should be offered, and the commentary to Principle 25, echoing the Human Rights Committee's General Comment, sets out a long list, including apologies, restitution, rehabilitation, financial or non-financial compensation and punitive sanctions (whether criminal or administrative, such as fines), as well as the prevention of harm through, for example, injunctions or guarantees of non-repetition.

One very simple way that businesses can be required to respect human rights is by placing a legal duty on them to do so. This is relatively rare, but there are jurisdictions which provide that their national human rights provisions apply to businesses. There are relatively few instances of individuals bringing claims using these opportunities. For example:

- Article 8(2) of the Constitution of South Africa: "A provision of the Bill of Rights binds a natural or a juristic person if, and to the extent that, it is applicable, taking into account the nature of the right and the nature of any duty imposed by the right."

- Article 20 (1) of the Constitution of Kenya: "The Bill of Rights applies to all and binds all State organs and persons" (and defines "person" as including "a company, association or other body of persons whether incorporated or unincorporated").

Alternatively, states can pass legislation which while not explicitly human rights legislation, can protect an individual's human rights. Privacy or data protection legislation, and equality or anti-discrimination legislation are two of the most common. For example:

- Data protection legislation (right to privacy), e.g. fine of £500,000 imposed upon Facebook by the UK's data protection authority following the Cambridge Analytica scandal, after Facebook allowed third party developers to access user information without sufficient consent.
- Equality or anti-discrimination legislation (rights to equality and non-discrimination), which could potentially be used against algorithms or automated decisionmaking which leads to discriminatory outcomes.

Operational principles

State-based judicial mechanisms

Principle 26 focuses on the first of the three grievance mechanisms, state-based judicial mechanisms, and how to ensure that these are accessible and effective.

Principle 26: States should take appropriate steps to ensure the effectiveness of domestic judicial mechanisms when addressing business-related human rights abuses, including considering ways to reduce legal, practical and other relevant barriers that could lead to a denial of access to remedy.

The commentary to Principle 26 looks at some general requirements which apply whatever kind of business or human rights impact is being considered. The focus of Principle 26 is on ensuring that judicial mechanisms are accessible and effective in addressing human rights abuses caused by businesses. General requirements include:

- The provision of justice should not prevented by corruption of the judicial process;
- Courts should be independent of economic or political pressures from other State agents and from business actors;
- The legitimate and peaceful activities of human rights defenders should not be obstructed.

There are a number of potential legal barriers which might be particularly relevant when it comes to human rights impacts caused by the tech sector. The first would be the absence of legislation which provides for individuals to bring claims against companies for particular types of adverse impacts. The second would be the existence of relevant legislation, but with provisions which mean it doesn't apply to the private sector, or to all types of businesses, or which can only be enforced by a regulatory body, rather than individual victims. The third would be discriminatory laws or rules which limit the ability of certain groups to bring cases, such as non-citizens or children.

Apart from legal barriers there are also some practical and procedural barriers which might be particularly relevant when it comes to human rights impacts caused by the tech sector. For example:

- High costs involved in bringing a claim and/or a lack of support through legal aid or litigation insurance
- Difficulties in obtaining legal representation
- Difficulties in bringing class actions, collective action claims or representative action by not-for-profit bodies, organisations or associations, when appropriate (such as an incident affecting the personal data of a large group of individuals)
- A lack of understanding of business and human rights, or the tech sector specifically, among relevant actors, including the judiciary

State-based judicial mechanisms are the bedrock of effective remedy but their effectiveness is based on a number of factors outlined above. Rights holders bringing claims against business will more often the not have fewer financial resources, and less access and expertise to judicial mechanisms and this can discourage claims being made or render claims useless. Judicial mechanisms need to remain impartial, independent and accessible particularly to disadvantaged and vulnerable groups. Furthermore, a strong judicial system which acts in rights holders best interests will encourage businesses to prevent human rights risks occurring in the first place.

Useful resources include:

- **Improving accountability and access to remedy for victims of business-related human rights abuse**: This report offers a set of practical resources which states can draw upon with a view to progressively and systematically improving their implementation of Pillar III of the UNGPs, including (i) a model terms of reference that can be used to review the effectiveness of domestic legal systems, (ii) an annex setting out a list of practical steps for

States to consider, arranged by themes ("policy objectives") relating to both procedural and substantive aspects of access to remedy.[45]

- **Illustrative examples for guidance to improve corporate accountability and access to judicial remedy for business-related human rights abuse**: A paper containing illustrative examples of methods that States have used and steps that States have taken in practice that are relevant to the different policy objectives and which have the potential to improve access to remedy in cases of business-related human rights abuses.[46]
- **The relevance of human rights due diligence to determinations of corporate liability**: A report that explores the relationship between human rights due diligence (as described in the UNGPs) and determinations of corporate liability under national law for adverse human rights impacts arising from or connected with business activities.[47]

State-based non-judicial grievance mechanisms

Principle 27 focuses on the second of the three grievance mechanisms, state-based non judicial grievance mechanisms and how to ensure that these are accessible and effective.

Principle 27: States should provide effective and appropriate non-judicial grievance mechanisms, alongside judicial mechanisms, as part of a comprehensive State-based system for the remedy of business-related human rights abuse.

Judicial mechanisms may not always be appropriate. Administrative, legislative and other non-judicial mechanisms play an essential role in complementing and supplementing judicial mechanisms.

Even where judicial systems are effective and well-resourced, they cannot carry the burden of addressing all alleged abuses; judicial remedy is not always required;

[45] UN Human Rights Council, Improving accountability and access to remedy for victims of business-related human rights abuse: Report of the United Nations High Commissioner for Human Rights, UN Doc. A/HRC/32/19, 10 May 2016, available at: https://www.ohchr.org/Documents/Issues/Business/DomesticLawRemedies/A_HRC_32_19_AEV.pdf

[46] UN Office of the High Commissioner for Human Rights, The OHCHR Accountability and Remedy Project: Illustrative examples for guidance to improve corporate accountability and access to judicial remedy for business-related human rights abuse, 5 July 2016, available at: https://www.ohchr.org/Documents/Issues/Business/DomesticLawRemedies/ARP_illustrative_examples_July2016.docx

[47] UN Human Rights Council, Improving accountability and access to remedy for victims of business-related human rights abuse: The relevance of human rights due diligence to determinations of corporate liability: Report of the United Nations High Commissioner for Human Rights, UN Doc. A/HRC/38/20/Add.2, 1 June 2018, available at: https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/20/Add.2

nor is it always the favoured approach for all claimants. Alternative, non-judicial mechanisms could include:

- National human rights institutions:
    - Could help facilitate class actions against tech companies where multiple users are affected, or bring cases on their behalf
    - Can support public awareness of how human rights can be adversely impacted by tech companies, and provide guidance on how to bring claims
    - Undertake research on the impacts on human rights by tech companies, helping bring evidence of potential abuses to light
- Ombudsperson offices
- Data protection authorities

Whilst being useful at filling gaps in judicial mechanisms, non-judicial mechanisms can pose difficulties. They may not be able to deal with large and complex cases, especially if they have a lack of investigatory power or if cases involve extraterritoriality. Even in smaller cases they may not be effective in the eyes of rights-holders due to a lack of independence from government, as many mechanisms will be government agencies. However, these mechanisms can be much more accessible financially, with many being free to use.

Useful resources include:

- **Improving accountability and access to remedy for victims of business-related human rights abuse through State-based non-judicial mechanisms**: In the report, the OHCHR explains the scope of the work involved and the approach taken by OHCHR, and makes general observations about the role of State-based non-judicial mechanisms in achieving accountability and access to remedy in business and human rights cases.[48]
- **State-based non-judicial mechanisms for accountability and remedy for business-related human rights abuses: Supporting actors or lead players?**: This paper identifies a number of legal, structural, practical and policy challenges. It also provides illustrative examples responding to the effectiveness criteria in Principle 31.[49]

[48] UN Human Rights Council, Improving accountability and access to remedy for victims of business-related human rights abuse: The relevance of human rights due diligence to determinations of corporate liability: Report of the United Nations High Commissioner for Human Rights, UN Doc. A/HRC/38/20/, 14 May 2018, available at: https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/20

[49] UN Office of the High Commissioner for Human Rights, Accountability and Remedy Project Part II: State-based non-judicial mechanisms, State-based non-judicial mechanisms for accountability and remedy for business-related human rights abuses: Supporting actors or lead players?, 2 November 2017, available at:

Non-state based grievance mechanisms

Principles 28, and 29 and 30 focuses on the third of the three grievance mechanisms - non-state-based grievance mechanisms - and how to ensure that these are accessible and effective.

Principle 28: States should consider ways to facilitate access to effective non-State-based grievance mechanisms dealing with business-related human rights harms.

Examples of non-state-based grievance mechanisms include:

- Those administered by a business enterprise alone or with stakeholders, by an industry association or a multi-stakeholder group;
- Regional and international human rights bodies.

Facilitating access from states could include information sharing and awareness-raising, as well as financial support.

Principle 29: To make it possible for grievances to be addressed early and remediated directly, business enterprises should establish or participate in effective operational-level grievance mechanisms for individuals and communities who may be adversely impacted.

Principle 29 calls on businesses to establish or participate in grievance mechanisms. There are particular advantages to businesses developing their own grievance mechanisms. They support the identification of adverse human rights impacts as a part of an enterprise's ongoing human rights due diligence, by providing a channel for those directly impacted by the enterprise's operations to raise concerns. They also make it possible for grievances, once identified, to be addressed and for adverse impacts to be remediated early and directly by the business enterprise, thereby preventing harms from compounding and grievances from escalating.

Grievance mechanisms are also a complement to stakeholder engagement as companies with effective procedures (as well as effective human rights due diligence procedures) will have most likely engaged with stakeholders during developing the grievance process or when grievances have occurred.

Examples of grievance mechanisms offered by tech companies include:

---

https://www.ohchr.org/Documents/Issues/Business/DomesticLawRemedies/ARPII_%20DiscussionpaperonPhase2forUNForum_FINAL.pdf

- Online platforms providing users the opportunity to challenge removals of content or the suspension of accounts
- Facebook's Oversight Board to independently review content moderation decisions
- Tech companies providing compensation for data breaches, or the misuse of personal data

Useful resources include:

- **Improving accountability and access to remedy in cases of business involvement in human rights abuses**: A paper that provides a preliminary assessment of current practices and challenges with respect to the use of non-State-based grievance mechanisms as a way of enhancing access to remedy in cases of business-related harm.[50]
- **Non-state based non-judicial grievance mechanisms (NSBGM): An exploratory analysis**: A paper highlighting the concept and existing practices of non-state based non-judicial mechanisms.[51]
- **How to Appeal?**: A guide developed by Online Censorship that guide users on how to appeal the decisions undertaken by Facebook, Twitter, Youtube, Instagram, Flickr and Medium.[52]
- **A Rights-Respecting Model of Online Content Regulation by Platforms**: A paper developed by Global Partners Digital proposing a model of online content regulation by platforms. On pages 23 and 24 it proposes the establishment of a grievance and remedial mechanism, allowing users to challenge decisions made to remove content (or suspend accounts), and to obtain an effective remedy where successful.[53]
- **Forgotten Pillar: The Telco Remedy Plan**: A guide to assist telcos to implement both the procedural aspects of remedy, such as safe and accessible grievance mechanisms, and the substantive aspects, which may be as simple as an explanation and commitment to non-repetition.[54]

---

[50] UN Office of the High Commissioner for Human Rights, OHCHR Accountability and Remedy Project: Improving accountability and access to remedy in cases of business involvement in human rights abuses: Phase III: Enhancing the effectiveness of non-State based grievance mechanisms Scope and Programme of Work, 1 November 2018, available at: https://www.ohchr.org/Documents/Issues/Business/ARP/ARPIII-PoW.pdf

[51] University of Manchester, Alliance Manchester Business School, Non-state based non-judicial grievance mechanisms (NSBGM): An exploratory analysis, 13 July 2018, available at: https://www.ohchr.org/Documents/Issues/Business/ARP/ManchesterStudy.pdf

[52] Online Censorship, How to Appeal, available at: https://onlinecensorship.org/resources/how-to-appeal

[53] Global Partners Digital, A Rights-Respecting Model of Content Regulation by Platforms, May 2018, available at: https://www.gp-digital.org/wp-content/uploads/2018/05/A-rights-respecting-model-of-online-content-regulation-by-platforms.pdf

[54] Acces Now, Forgotten Pillar: The Telco Remedy Plan, May 2013, available at: https://www.accessnow.org/cms/assets/uploads/archive/Telco%20Remedy%20Plan.pdf

Principle 30: Industry, multi-stakeholder and other collaborative initiatives that are based on respect for human rights-related standards should ensure that effective grievance mechanisms are available.

An effective initiative, whether industry, multi-stakeholder or collaborative, should have effective grievance and remedy processes. While they should be open to individuals to use, they shouldn't remove or restrict their ability to use grievance mechanisms offered by individual companies. Examples of these types of grievance mechanisms in the tech sector are, however, rare.

Effectiveness criteria for non-state based grievance mechanisms

Principle 31: In order to ensure their effectiveness, non-judicial grievance mechanisms, both State-based and non-State-based, should be:

(a) Legitimate: enabling trust from the stakeholder groups for whose use they are intended, and being accountable for the fair conduct of grievance processes;

(b) Accessible: being known to all stakeholder groups for whose use they are intended, and providing adequate assistance for those who may face particular barriers to access;

(c) Predictable: providing a clear and known procedure with an indicative time frame for each stage, and clarity on the types of process and outcome available and means of monitoring implementation;

(d) Equitable: seeking to ensure that aggrieved parties have reasonable access to sources of information, advice and expertise necessary to engage in a grievance process on fair, informed and respectful terms;

(e) Transparent: keeping parties to a grievance informed about its progress, and providing sufficient information about the mechanism's performance to build confidence in its effectiveness and meet any public interest at stake;

(f) Rights-compatible: ensuring that outcomes and remedies accord with internationally recognised human rights;

(g) A source of continuous learning: drawing on relevant measures to identify lessons for improving the mechanism and preventing future grievances and harms;

Operational-level mechanisms should also be:

(h) Based on engagement and dialogue: consulting the stakeholder groups for whose use they are intended on their design and performance, and focusing on dialogue as the means to address and resolve grievances.

UNGP 31(a) Legitimate: The process needs to be one which is trustworthy, ensuring that the parties to a grievance process cannot interfere with its fair conduct. Given the broad range of potentially affected users, institutions administering the mechanisms should pay particular attention to ensuring that design of mechanisms involves meaningful participation by affected stakeholders and user groups, with particular attention paid to the needs of vulnerable and marginalised groups.

UNGP 31(b) Accessible: Tech companies have an advantage when it comes to ensuring that grievance mechanisms are accessible, since the products and platforms that they develop and use can themselves be used to raise awareness of the grievance mechanisms. A particular barrier as regards accessibility of grievance mechanisms in the tech sector, however, is language. Many tech companies are global in nature and can adversely impact the rights of a wide range of people who speak many different languages. Ensuring that grievance mechanisms can impart and receive information in appropriate languages is crucial so that users are able to engage grievance processes at all.

UNGP 31(c) Predictable: The grievance and remediation process should be made public, as well as promoted so that people are encouraged to use it. There should be, at a minimum:

- Clarity, in plain language, on the types of processes that are available
- Clear, known procedures (including any procedural rules) with indicative time frames for each stage (as well as processes in place to ensure that those time frames are respected)
- Clarity on the types of outcomes that are available
- Clarity on the means of monitoring and implementation of those outcomes

UNGP 31(d) Equitable: Aggrieved parties should have access to the necessary expertise, advice and information so that they can engage in grievance mechanisms on fair terms. This means that any policies or rules relating to the grievance mechanism should be clear and understandable. If necessary, additional financial or other resources should be made available. The grievance mechanism should enable group actions to be made, or for grievances to be made by representatives of rights-holders if necessary.

UNGP 31(e) Transparent: Institutions administering the mechanisms should look to publish information on current and historic grievances that have been brought. This could be done through annual reports which detail information on the number of grievances raised, how these were dealt with, and what the outcomes were. Confidentiality is essential, particularly so as to ensure that human rights such as the right to privacy are not put at risk. When a user has raised a grievance against a company they should be regularly informed of the progress of their complaint.

UNGP 31(f) Rights compatible: Institutions administering the mechanisms should assess the level of satisfaction with outcomes from those raising complaints, as well as broader stakeholder groups, ensuring that outcomes are compatible with international human rights standards. This requires remedies not to infringe human rights themselves. This could mean, for example, ensuring that the identities of those raising grievances are not made public, so as to respect the right to privacy.

UNGP 31(g) A source of continuous learning: Institutions administering the mechanisms should regularly analyse the frequency, patterns and causes of grievances, as well as levels of satisfaction of those submitting grievances. Such information can be used to help ensure that the grievance process is continually developed to meet the needs of users and would-be-users. This information should also be used to assess and improve policies, procedures and practices of the entity responsible for the harm so as to prevent future grievances. Human rights abuses by tech companies can often be quite geographically specific due to certain states' regulations, attitude or actions when it comes to digital rights so it is sensible for companies to assess trends to rectify these issues in particular states.

UNGP 31(h) Operational-level mechanisms should also be based on engagement and dialogue: As noted under Principle 31(a), institutions administering the mechanisms should pay particular attention to ensuring that design of mechanisms involves meaningful participation by affected stakeholders and user groups, with particular attention paid to the needs of vulnerable and marginalised groups. Given their global reach, tech companies should pay particular attention to proactive outreach programmes to raise awareness about stakeholder rights, the grievance mechanism, and how to use it. These can be done through the products and platforms that the tech companies develop. Tech companies should also consider involving independent third parties in appropriate cases.

Useful resources include:

- **Accountability and Remedy Project Part III: Non-State-based grievance mechanisms: Enhancing effectiveness of non-State-based grievance mechanisms in cases of business-related human rights abuse:** An OHCHR discussion paper, the annex of which sets out some illustrative examples of

features of non-state-based grievance mechanisms potentially relevant to the implementation of Principle 31 of the UNGPs.[55]

[55] UN Office of the High Commissioner for Human Rights, OHCHR Accountability and Remedy Project: Part III: Non-State-based grievance mechanisms: Enhancing effectiveness of non-State-based grievance mechanisms in cases of business-related human rights abuse: Discussion Paper, 19 November 2019, available at: https://www.ohchr.org/Documents/Issues/Business/ARP/ARPIII_Discussion_Paper_Nov2019.pdf.