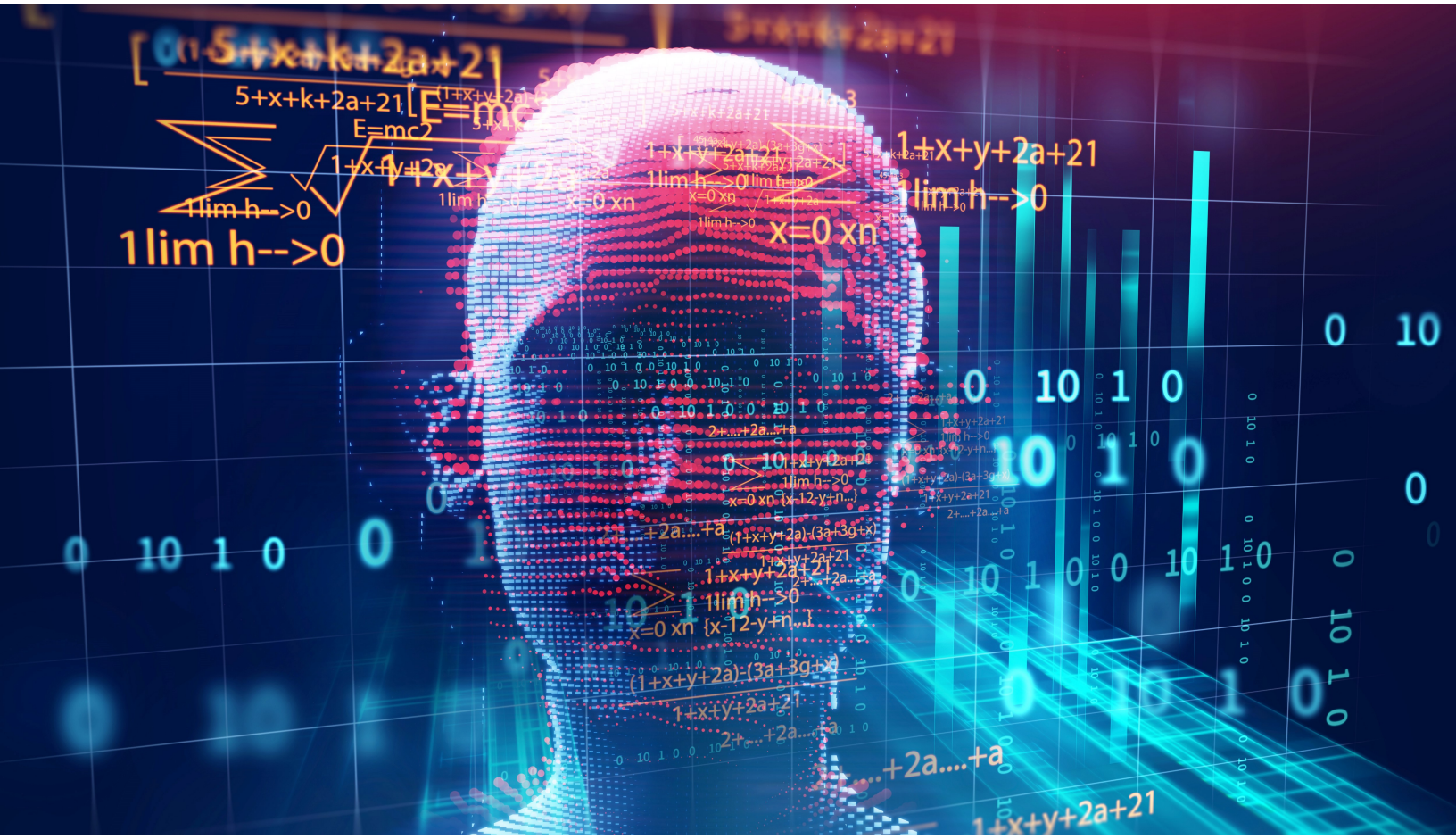
WORKING PAPER

AUGUST 2018

# Artificial Intelligence: A Rights-Based Blueprint for Business

## Paper 1: Why a Rights-Based Approach?



BSR
The Business of a Better World

# About This Report

This report was written by Dunstan Allison-Hope (Managing Director, BSR) and Mark Hodge, an independent business and human rights expert.

Artificial Intelligence (AI) technologies—and the big data business models underpinning them—are disrupting how we live, interact, work, do business, and govern. The economic, social, and environmental benefits of AI could be significant. But as evidence mounts about potential negative consequences for society and individuals, we urgently need a robust view of what responsible conduct looks like and a vision for how markets and governance mechanisms can guide the right behaviors.

We believe that the speed, complexity, and extensive reach of AI requires an approach to responsible practice that is rights-based. In three papers we draw upon approaches and lessons learned from the field of business and human rights to describe a blueprint for responsible business practice both within and beyond the technology sector.

Deliberate investment in rights-based approaches is urgently needed to avoid two risks: First, that new technologies, capabilities, and business models are unleashed into the world that cause significant harm to the rights to which all human beings are inherently entitled; and second, that a once-in-a-generation opportunity to harness massive advances in technology for the public good is missed.

This is the first of three working papers intended to develop and test new business policies and practices aimed at establishing a sustainable social license to operate for new AI technologies that are capable of creating long-term sustainable value for all stakeholders.

- In this first paper we outline 10 beliefs—built on the internationally agreed-upon foundations of the business and human rights field—to govern and guide the use of AI. We draw heavily on the *United Nations Guiding Principles on Business and Human Rights (*UNGPs), the foundational and internationally endorsed road map for addressing business human rights impacts on people.

- The second paper argues for attention to be paid to the AI value chain and demonstrates that the positive and negative human rights impacts associated with AI are directly relevant for companies beyond the technology sector.

- The third paper explores what tools, methodologies, and guidance exist or will need to exist to operationalize business respect for human rights in the context of AI development and use.

These three papers have been based on a mixture of desk-based research and direct experience by the authors engaging with business on human rights due diligence. They are positioned as “working papers” to stimulate discussion and influence the ongoing debate. The authors welcome feedback, comment, and dialogue on the papers, and we look forward to working with others to shape the next iteration of these ideas.

Please direct comments or questions to web@bsr.org.

## ACKNOWLEDGMENTS

The authors wish to thank Elisabeth Best, Hannah Darnton, Michael Karimian, Michaela Lee, Peter Nestor, Moira Oliver, Jacob Park, and Michael Rohwer for their review, insights, and guidance. Any errors that remain are those of the authors.

## DISCLAIMER



## SUGGESTED CITATION

Allison-Hope, Dunstan and Hodge, Mark 2018, “Artificial Intelligence: A Rights-Based Blueprint for Business,” BSR, San Francisco.

# Contents

# Introduction

Artificial Intelligence (AI) can be defined as intelligence exhibited by machines. It includes both "machine learning" (an approach to achieve AI), which uses algorithms to parse data, learn from it, and then make a determination or prediction, and "deep learning" (a technique for implementing machine learning), which is inspired by understanding the biology of our brains.

AI is advancing rapidly, thanks to ever-more-powerful computing, massive growth in the availability of digital data, and increasingly sophisticated algorithms. These advances bring enormous opportunities to address big social challenges, such as improved health diagnostics, self-driving vehicles that improve road safety, and enhanced fraud prevention, to name just three. AI also brings social risks, including new forms of discrimination arising from algorithmic bias, labor impacts arising from the displacement of workers by machines, increased potential of surveillance by employers and the State using tracking devices and facial recognition tools, and new risks to child rights as the volume of data collected about children increases substantially.

Various ethics-based approaches to the responsible development and deployment of AI have emerged over the past two years that cover items such as privacy, surveillance, discrimination, bias, unintended consequences, and misuse by bad actors.[1] These have all made a tremendously positive contribution to the debate about the future of AI. We believe that rights-based approaches offer a robust framework for the responsible development and use of AI and should form an essential part of business policy and practice.

There is a growing body of opinion leaders, from civil society, technology groups, business, government, and professional bodies, arguing that the protection, respect, and fulfillment of human rights should be at the core of the vision for AI. Civil society groups and worker organizations have been exploring the social, ethical, and human rights implications of data-driven technologies. A national human rights commission has launched a project that will seek to ensure human rights are prioritized in the design and regulation of new technologies. Professional associations reference human rights in their efforts to create new principles and codes.[2] In May 2018, a coalition of human rights organizations and technology groups signed the Toronto Declaration, calling on governments and technology companies to ensure that machine learning systems do not undermine equality and the right to nondiscrimination.[3]

As a contribution to this nascent consensus, we use this paper to describe why the principles and core concepts within the business and human rights field offer a compelling baseline from which to define and implement responsible business practice. We do this through the articulation of 10 beliefs that are grounded in the *UN Guiding Principles on Business and Human Rights* (UNGPs).[4] Together, these beliefs elaborate on three cross-cutting messages about a rights-based approach and the value of business and human rights frameworks:

---

[1] Examples include the Ethical OS Toolkit, Future of Life Principles, and the ITI Policy Principles on AI

[2] For example, see the IEEE Global Initiative on the Ethics of Autonomous and Intelligent Systems and the ITI AI Policy Principles

[3] https://www.accessnow.org/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems/

[4] https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

» **Holistic, detailed, and predictable:** A rights-based approach defines business responsibility and accountability with reference to a list of precisely crafted standards that have been agreed upon by societies and the international community, offering certainty and clarity to all actors. Actions are judged against these clear standards, not against our (or someone else's) individual preferences or ethos. A rights-based approach also emphasizes the interdependence and simultaneously equal importance of the civil, political, economic, cultural, and social circumstances that allow every human to thrive—there are no blind spots, and there is no unintentional undercutting of one aspect of human well-being over another.

» **Individual accountability alongside collective problem solving**: The business and human rights field takes a systems-thinking approach to embedding responsible business practices and behaviors. While each individual actor (whether businesses or governments) must meet clear standards of conduct, the networked and global nature of 21st century society means that we are all part of the problem and all part of the solution.

» **Key components of responsible business conduct are already clarified:** The business and human rights field has already established a road map for what companies need to do to operate with respect for human rights. Companies and their stakeholders in sectors as diverse as apparel, agriculture, mining, and electronics are already developing and applying policies, tools, and approaches to address and mitigate the actual and potential impacts that their activities create.

## 10 Beliefs
## for the responsible development and use of AI

| | |
|---|---|
| 1 | The development and use of AI should be done in ways that respect all clearly articulated and internationally agreed-upon human rights. |
| 2 | We need both governance and technical solutions for the responsible development and use of AI. |
| 3 | Special attention should be paid to the State-business nexus, especially the use of private sector AI solutions in public service delivery. |
| 4 | All actors in all industries across the AI value chain have responsibilities—including those buying and using AI solutions outside of the technology sector. |
| 5 | Responsible business conduct is about the business models and strategies used by companies to take AI to market, not just the risk and merits of specific AI technologies. |
| 6 | AI brings new and previously unforeseen human rights challenges, and the onus is on businesses to proactively "know and show" how they address the actual or potential adverse impacts of AI. |
| 7 | AI also brings challenges like those previously experienced in other industries, and we can learn from them. |
| 8 | User-centered design should address the experiences and views of people, especially vulnerable populations, who may be negatively impacted by new technologies. |
| 9 | Those whose human rights have been violated—however unintentionally—by the deployment of AI solutions should have access to remedy. |
| 10 | We should stretch business and human rights methodologies to suit the nature and pace of AI development and deployment. For example, we should explore rights-based approaches for maximizing the positive impacts of AI, and experiment with the use of foresight methodologies in human rights due diligence. |

## 1: The development and use of AI should be done in ways that respect all clearly articulated and internationally agreed-upon human rights.

The business and human rights field has established that any company in any industry can impact—positively or negatively—any of the internationally agreed-upon human rights. For example, extractives companies can have impacts beyond the health and safety of workers or the well-being and security of local communities, such as forced labor used in the construction phase of mining operations or sexual exploitation of women due to the presence of male workers living around a mining site.

Similarly, those seeking to develop and use AI responsibly should start with the assumption that their activities, products, and services can impact any of the 30 human rights set out in the Universal Declaration of Human Rights (UDHR). Although certain rights will obviously merit immediate and urgent focus, such as the right to privacy or nondiscrimination, it is a mistaken and bad risk management strategy (for business and society) to think narrowly about impacts.

A well-discussed challenge where a rights-based approach may have helped is the impact of social media on the right to political participation. A wide-ranging rights-based analysis might have surfaced the risk that targeted campaigning and the deliberate spread of misinformation could lead to abuses of the right to political participation and the right to take part in elections without unreasonable infringements.

There are four reasons why international human rights standards and the use of rights-based approaches offer a robust, practical, and globally applicable normative framework upon which a responsible approach to the development and deployment of AI should be based.

- By providing a common standard of achievement for all peoples and a "long list" of rights—as diverse as land rights, freedom of expression, and a decent standard of living—international human rights standards enable us to avoid blind spots and help ensure that a full range of potential positive and negative impacts are systematically considered.

- The use of international standards as a baseline to define responsible business conduct befits a globalized economy and a moment in history in which information, exchange, and business takes place online and across borders. Companies should follow these standards, regardless of national laws, and doing so is fundamental to recognizing that everyone's dignity and rights should be upheld and respected regardless of ethnicity, race, gender, nationality, age, or sexual preference.

- A crucial doctrine of the business and human rights field is that a company cannot "offset" negative impacts on people by pointing to positive impacts elsewhere. This is one way in which the business and human rights field raises the bar on other schools of corporate responsibility whereby (for example) a mining company might claim to offset increases in respiratory disease among children by providing philanthropic donations to a global health NGO. The same doctrine should apply to AI—no matter the scale of positive potential, there should be constant innovation to prevent, mitigate, and avoid, rather than accept, negative impacts.

- Human rights provide business leaders, engineers, and developers with a globally applicable set of standards agreed by the international community that provide clarity and certainty for how to develop and deploy AI in ways that support the inherent dignity and inalienable rights of all human beings.

## 2: We need both governance and technical solutions for the responsible development and use of AI.

A decade ago, the architect of the UNGPs articulated that "the root cause of the business and human rights predicament today lies in the governance gaps created by globalization—between the scope and impact of economic forces and actors, and the capacity of societies to manage their adverse consequences."[5] The parallels to the current nature and speed of technological development are clear, and we may already be seeing a lack of capacity of society to manage the adverse consequences of disruption. As a way forward, the UNGPs offer an elegant solution.

On the one hand, the UNGPs focus on the various ways in which States—individually and in unison—can apply a smart mix of measures to address governance gaps. Such measures might involve the use of myriad tools available to governments, including regulation, guidelines, public procurement, corporate reporting requirements, investment contracts, and development finance.

On the other hand, the UNGPs also focus on the need to move beyond business as usual to embed respect for human rights into core business activities so that practical measures are put in place to prevent and mitigate adverse impacts. This can often lead to "technical" and operational solutions to entrenched problems.

When considering the challenge of ensuring responsible development and use of AI, this duality of governance and technical problem-solving should be embraced. For example, government guidelines, regulation, and enforcement will be needed to ensure that businesses remain accountable for discrimination in job advertising, hiring, firing, and promotion decisions when these are guided by algorithms. At the same time, we should welcome, encourage, and reward technical solutions that weed out discrimination in such tools and indeed the development of tools that are an improvement on human bias. The use of data sheets to address bias in training data-sets is one such example.[6]

An implication of this proposition is that arguing between libertarian individualism and State paternalism is likely to distract from finding practical solutions, prevent international collaboration, and ignore the critical role of civil society to shine a light on abuse and advocate for clearer standards. A human rights approach and a focus on what works for the vulnerable can be applied beyond such political and ideological preferences.

Certainly, State action can build momentum and provoke laggards to act, as we have seen in the realm of human trafficking and modern slavery. At the same time, there are plenty of examples in the business and human rights field in which industry associations (such as the Responsible Business Alliance or the International Council on Mining and Metals) and multistakeholder initiatives (such as the Fair Labor Association, the Global Network Initiative, or the Centre for Sport and Human Rights) have addressed human rights issues embedded in the global economy in ways that go beyond where States have been willing or able to, and where companies and stakeholders have achieved more together than they could alone.

---

[5] https://documents-dds-ny.un.org/doc/UNDOC/GEN/G08/128/61/PDF/G0812861.pdf?OpenElement

[6] https://arxiv.org/abs/1803.09010

## 3: Special attention should be paid to the State-business nexus, especially the use of private sector AI solutions in public service delivery.

A large portion of the State Duty to Protect pillar of the UNGPs addresses the "State-business nexus." The situations that the relevant UNGP principles and commentary cover include: when the State provides export credit and investment insurance; when the State legislates for or enters into contracts with the private sector as part of public service delivery; and when the State procures products or services from the private sector. The essential message is that in such situations States cannot contract out their primary duties to protect, respect, and fulfill human rights obligations.

The deployment of privately developed big data and AI solutions to support the delivery of public services is already a central focus of attention for civil society and leading experts, such as in the areas of law enforcement, criminal sentencing, smart cities, access to health, allocation of funding in the public education system, intelligence gathering, and defense. When reading case studies and news stories about the negative human rights consequences on ordinary people (often the most vulnerable in society), it appears that States are insufficiently aware of the potential downsides for their citizens of deploying new technologies.

The UNGPs posit that States should require that enterprises implement human rights due diligence when services are contracted or outsourced to them, and "ensure that they [States] can effectively oversee the enterprises' activities, including through the provision of adequate independent monitoring and accountability mechanisms."[7] This requirement—and the related field of practice and expertise from the business and human rights field—clearly offer one clear anchor point to address the use of AI technologies by the State.

It will also be critical for companies to undertake their own due diligence regarding how their AI solutions are deployed by States and avoid scenarios where the risk of adverse impacts on human rights are most severe.

The importance of this proposition is reinforced by the fact that these scenarios often involve essential public services provided to the most vulnerable populations. However, it would be wrong to suggest that this proposition is only about risk mitigation. The use of new technologies to collect and process vast amounts of data in poorly served areas (such as in the Global South) and the use of AI to inform poverty alleviation, humanitarian interventions, and better access to public services for the most vulnerable clearly represent opportunities to enhance the realization of human rights for those who need it the most.

[7] Principle 5, https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

## A CRITICAL SIDE NOTE:

## STATE ACCOUNTABILITY AND THE WIDER HUMAN RIGHTS MOVEMENT

We believe that applying business and human rights approaches to the private sector's role and responsibility when developing and deploying AI is a compelling and practicable proposal. At the same time, sustaining the wider human rights movement—especially in the context of multilateral institutions, State sanctions, and the continued work of human rights defenders, victim's groups, and civil society—is critical.

States in all regions of the world are seeking to use AI as part of public service delivery. In a context where international human rights are well embedded nationally and offer meaningful protections of vulnerable groups, the risks will be significantly reduced. But where gaps or regressive laws are in place, AI can amplify vulnerability. For example, if certain segments of the populations are discriminated against based on such characteristics as ethnicity, race, or sexual preference, then AI is highly likely to enable States to do this in more systematic ways. The same can be said for other human rights issues such as privacy, right to information, and freedom of expression.

Further, States will increasingly look to AI to strengthen their economic and geopolitical position in the world, such as by using new technologies in their intelligence efforts and in the context of military spending. This opens the door for a range of nefarious uses, the likes of which will not be made transparent.

Such realities make local and international efforts to secure protections and legal recourse ever more urgent. As we note in our third paper, responsible businesses should play their part by adopting ongoing due diligence regarding public sector customers, monitoring of end-use, and engaging thoughtfully in collective action and public policy. But this should not reduce focus and investment in the work of the existing human rights movement.

## 4: All actors in all industries across the AI value chain have responsibilities—including those buying and using AI solutions outside of the technology sector.

A central challenge that the business and human rights field has sought to address is how to conceptualize the responsibility of a given enterprise when adverse impacts on human rights occur in the value chain of that company, even where that enterprise is not necessarily causing the adverse impacts to occur.

This is a critical question when we consider the potential impacts of the use of technology products and services. Is the developer and seller of solutions (for example, data-driven insights about social media users) responsible for the misuse of those insights by a customer? If a technology company sells facial recognition technology that discriminates because it was trained on biased data-sets, is that the fault of the selling company, or of the customer that has integrated that tool into its offerings to provide security services to airports? What if a firm designs and sells geo-location technologies to a mobile phone company, and this functionality is then abused by an authoritarian government to target and imprison protesters?

The UNGPs offer an elegantly simple answer. Every enterprise in the value chain has responsibility to exercise due diligence and use their leverage to avoid negative outcomes for people, regardless of whether that business is designing, developing, marketing, selling, buying, using, or somehow benefiting from that technology. The UNGPs do differentiate between what different actors are expected to do based on the degree of their involvement in the adverse impact, but the underlying premise does not waiver.

The implications of this way of thinking about responsibility in complex, multi-tiered, global value chains for data-driven and AI technologies include:

» Non-technology companies buying and using data-driven and AI solutions should exercise diligence in what they are buying. They should procure responsibly.

» Companies selling technology solutions should be responsible product (and service) stewards, even after a product has been sold to a third party.

» Other private actors in the technology value chain that have responsibilities, such as research and development institutions and investors, should be exercising human rights due diligence.

Addressing the question of who is ultimately responsible when algorithms and decision-making may not be explainable is undoubtedly challenging, but the UNGPs would seem to suggest a strong element of shared responsibility and accountability. It also suggests that companies outside the technology industry should be actively engaged in human rights due diligence in the context of AI, which we address in the second paper of this series.

## 5: Responsible business conduct is about the business models and strategies used by companies to take AI to market, not just the risk and merits of specific AI technologies.

The business and human rights field requires that responsible businesses address the actual and potential impacts of business models, competitive strategies, and the ways in which products and services are developed, marketed, and sold. In short, it is a holistic perspective concerned with the myriad ways that any given company creates value. This includes scrutiny of the public policy, lobbying, and legal strategies that a company deploys on issues that could have a substantial impact on the realization of human rights.

In most cases, reaching simplistic conclusions that a given technology is either "good" or "bad" is counter-productive and could do more harm than good. For example, the idea of providing access to credit for economically struggling families so that they can enter the property market and buy homes is certainly not bad in its own right and can bring significant human rights benefits. However, the product development, short-term profiteering, and opportunism that led to the late 2000's subprime mortgage collapse and left millions of people homeless or unemployed clearly had negative consequences for human rights. Another example is that weapons can be legitimately used to protect citizens and maintain security, but when weapons are sold to actors seeking to intimidate and commit crimes against their own population, then the same technologies can have serious impacts on human rights. So, the focus becomes as much about how, and to whom, arms are traded, and less (with some clear redlines) about the product or technology itself.

This broad understanding of responsibility is helpful when considering the responsible development and use of data-driven and AI technologies for the following reasons:

» It moves us beyond the philosophical discussions about whether new technologies are good or bad for human rights, and toward a focus on the careful use of technologies in varying contexts. This is especially critical due to the tendency in the media to sensationalize and scare-monger via an "AI evil vs. AI good" story-line.

» While engineers, developers, and data scientists can play a crucial role in responsible design of AI technologies and solutions, the business context, commercial incentives, and competitive pressures merit significant exploration too.

» The role of companies in influencing the legal, regulatory, and policy frameworks that shape the deployment of AI should be consistent with the commitments they make to respect human rights and be good corporate citizens.

## 6: AI brings new and previously unforeseen human rights challenges, and the onus is on businesses to proactively "know and show" how they address the actual or potential adverse impacts of AI.

The UNGPs establish that companies—independent of legal requirements or demands from customers, investors, consumers, users, and citizens—should be proactive in identifying ("knowing") and communicating ("showing") the potentially adverse human rights impacts of their business *and the steps they are taking to prevent and mitigate* those potentially adverse impacts. The UNGPs note that companies should have in place "a human rights due diligence process to identify, prevent, mitigate, and account for how they address their impacts on human rights."

This "know and show" aspect of the UNGPs is especially important in the context of AI given two factors:

- Expertise on AI remains highly specialized and concentrated inside engineering functions at a small number of the world's largest companies.

- In contrast to more established business models and technologies, the average person has a very limited understanding of how data, code, and algorithms work. This is perhaps the first time in human history that companies are deliberately bringing to market technologies and solutions where, sometimes by design, we do not fully understand how they are operating.

This disparity in knowledge and power can only be balanced by doing exactly what the UNGPs call for: that companies transparently and proactively demystify for affected rights-holders and wider society—the public, civil society organizations, policy-makers, and others—the actual and potential human rights impacts they believe may arise from the deployment of AI by their business.

Moreover, while there are clear human rights reasons for following the UN guidelines, there is also a business rationale for doing so—namely, that building trust and avoiding and mitigating risks to people can easily become legal, financial, reputational, and operational risks for companies.

## 7: AI also brings challenges like those previously experienced in other industries, and we can learn from them.

There are certainly unique characteristics to the so-called fourth industrial revolution or "second machine age" that will lead to risks to human rights manifesting in different ways. At the same time, a range of issues—such as discrimination; job losses due to labor-saving practices; product stewardship; and free, prior, and informed consent—have been challenges for many other industries. Beyond specific human rights impacts, the past few decades have seen diverse industries seek to embed policies, processes, management systems, and stakeholder engagement efforts that seek to mitigate the bad effects on people that a company could be involved with.

Attempts by other industries to address human rights issues that could offer insight to responsible development and use of AI include the following:

» Approaching road safety not just from the perspective of a user action (or error), but also with regard to automobile design—demonstrating that design can make a significant difference to the number of accidents and fatalities.

» Free, prior, and informed consent (FPIC) for indigenous peoples or in clinical trials in the pharmaceutical sectors—demonstrating that FPIC could have relevance in how we all, especially vulnerable groups, consent to the use of personal data in AI solutions, particularly in higher-risk scenarios.

» The engagement of civil engineering functions in efforts to address risks to communities of large infrastructure projects, such as cutting off access to cultural sites or making children's walks to schools more dangerous—demonstrating that engaging engineering functions early and often in developing effective mitigation actions can be effective.

» Taking steps to respect the right to health, for example via nutrition labeling in the food and beverage industry—demonstrating the role that companies can play in educating consumers about purchasing, product use, and consumption behaviors.

» Seeking to embed commitments and actions to uphold labor rights in value chains by mapping complexity, tracing commodities, embedding requirements in contracts, capacity-building, and identifying leverage points to address root causes—demonstrating the relevance of approaches that engage the entire value chain.

More opportunities can and should be found. However, our key message is that generating responsible and rights-respecting business practice in relation to AI need not start from a blank slate or proceed on a simplified conception that every issue related to new technologies are exceptional and brand new. One does not always need to reinvent the wheel in terms of norms and good practice—tactics and lessons from other industries can be borrowed and adapted.

## 8: User-centered design should address the experiences and views of people, especially vulnerable populations, who may be negatively impacted by new technologies.

The idea of including users in the design and testing of new products and services is not new. What is perhaps unusual is the notion that groups who may be negatively impacted by a technology should be involved in their design and methods of promotion and use.

However unusual at face value, engaging affected people is at the essence of the business and human rights agenda, and the golden thread of the UNGPs. Furthermore, leading firms in diverse industries have begun to embrace this practice as a moral and business imperative. Being listened to about the negative impacts that the actions of a third party have on us and then being genuinely involved in finding solutions and ways forward, is central to the experience of dignity. Further, a company that fails to pay attention to and address the concerns, frustrations, and bad experiences of people that it (however inadvertently) affects will experience some form of social, and even legal, backlash.

In any industry and operational context this is not always an easy proposition to fulfill. Issues include that the users or affected populations rarely experience harms and benefits in a uniform way, or that in many situations a company may lack direct relationships with affected people (for example, workers deep in the supply chain or consumers of a product that is being sold by a third-party retailer). Nonetheless, inventive approaches and methods exist in a range of industries that the developers and users of AI solutions could build on, including:

- Identifying legitimate representative or proxy organizations to engage with.
- Using technology to enable affected rights-holders to share grievances and experiences.
- Engaging directly with individuals who have suffered abuses in the past.
- Diversifying the workforce to involve viewpoints informed by a wide variety of life experiences and challenges in company discussions.

## 9: Those whose human rights have been violated—however unintentionally—by the deployment of AI solutions should have access to remedy.

The third pillar of the UNGPs establishes that access to remedy should be provided for victims of business-related abuses. However, the rapid development of AI raises three new challenges for securing access to remedy, which can be especially challenging when humans often can't cognitively understand how a decision is made by AI systems:

- Guaranteeing remedy when violations result from decisions made by machines and algorithms, rather than humans.
- Providing operational grievance mechanisms when there are hundreds of millions of rights-holders and billions of decisions.
- Safeguarding access to remedy when dozens of companies, rather than a single corporate actor, are linked to a human rights violation via the interaction of different AI-based products and services.

A growing body of thought exists on access to remedy in the context of AI, such as debates that have taken place on how to interpret the "right to explanation" for machine-based decision-making contained in the EU General Data Protection Regulation. However, further exploration of specific uses cases is needed, and (as described in Belief Eight), it is essential that civil society, rights-holders, and vulnerable populations benefit from new channels to participate meaningfully in the development of systems for how access to remedy can be secured in practice.

It is also important that by access to remedy fulfills the effectiveness criteria set out in Principle 31 of the UNGPs by being legitimate, accessible, predictable, equitable, and transparent.

## 10: We should stretch business and human rights methodologies to suit the nature and pace of AI development and deployment. For example, we should explore rights-based approaches for maximizing the positive impacts of AI and experiment with the use of foresight methodologies in human rights due diligence.

We are firm in the contention that the foundational concepts from the business and human rights field and the related business practices that have evolved in the past few decades, offer a rigorous foundation for creating a world in which big data and AI solutions address, rather than increase, vulnerability. However, there are certain characteristics of the current era of technological disruption that may require us to stretch the field itself.

Two issues that clearly merit attention are:

» **Human rights opportunities:** The possibility of using the underlying principles of human rights due diligence as a tool to drive and scrutinize the positive human rights potential of AI technologies. Is there an opportunity to create "rights-based innovation"?

» **Future-testing human rights due diligence:** The fact that we are deeply uncertain about how technologies will evolve in the years and decades ahead makes human rights due diligence of AI very challenging. It can be nearly impossible to credibly predict all the positive possibilities and negative implications for how we live, work, govern, and organize arising from the deployment of AI. Methodologies for implementing respect for human rights may need to integrate strategic tools such as strategic foresight, futures thinking, and scenario planning.

These and other considerations about how to undertake human rights due diligence in the context of AI are considered in the final of our three working papers.

# Looking Ahead

The recent turn to rights-based thinking to address the potential harms of disruptive technologies is welcome. We believe it can pave the way for maximizing the potential of AI to unlock humanity's inventiveness and creativity, improve lives, and increase the realization of human rights.

We believe that the business and human rights field provides the most intellectually robust and actionable framework to clarify and develop responsible business practice. We have offered 10 beliefs—based heavily on the foundations of the UN Guiding Principles on Business and Human Rights—to spark dialogue among leaders from business, civil society, and government on a path forward. We look forward to exploring the level of multistakeholder consensus (or disagreement) that might exist around these ideas, and to discussing how these propositions can be used to embed dignity for all into today's era of technology-based disruption.

Some questions to consider include:

» Do stakeholders from business, civil society, investors, government—and across diverse regions of the world—believe that our case offers a robust and valuable operating system for business?

» How might our case be further researched and developed in unison with existing initiatives focused on ethics or other frameworks?

» What business practice can we already point to that follows the approach described in this document? What can we learn from such efforts about how to shape and scale new behaviors?

This series of papers has been developed as "working papers" to stimulate discussion and influence the ongoing debate about the responsible use of AI by business. The authors welcome feedback, comment, and dialogue on the papers, and we look forward to working with others to shape the next iteration of these ideas.

Please direct comments or questions to web@bsr.org.

## About BSR

BSR is a global nonprofit organization that works with its network of more than 250 member companies and other partners to build a just and sustainable world. From its offices in Asia, Europe, and North America, BSR develops sustainable business strategies and solutions through consulting, research, and cross-sector collaboration. Visit www.bsr.org for more information about BSR's 25 years of leadership in sustainability.