



BSR®

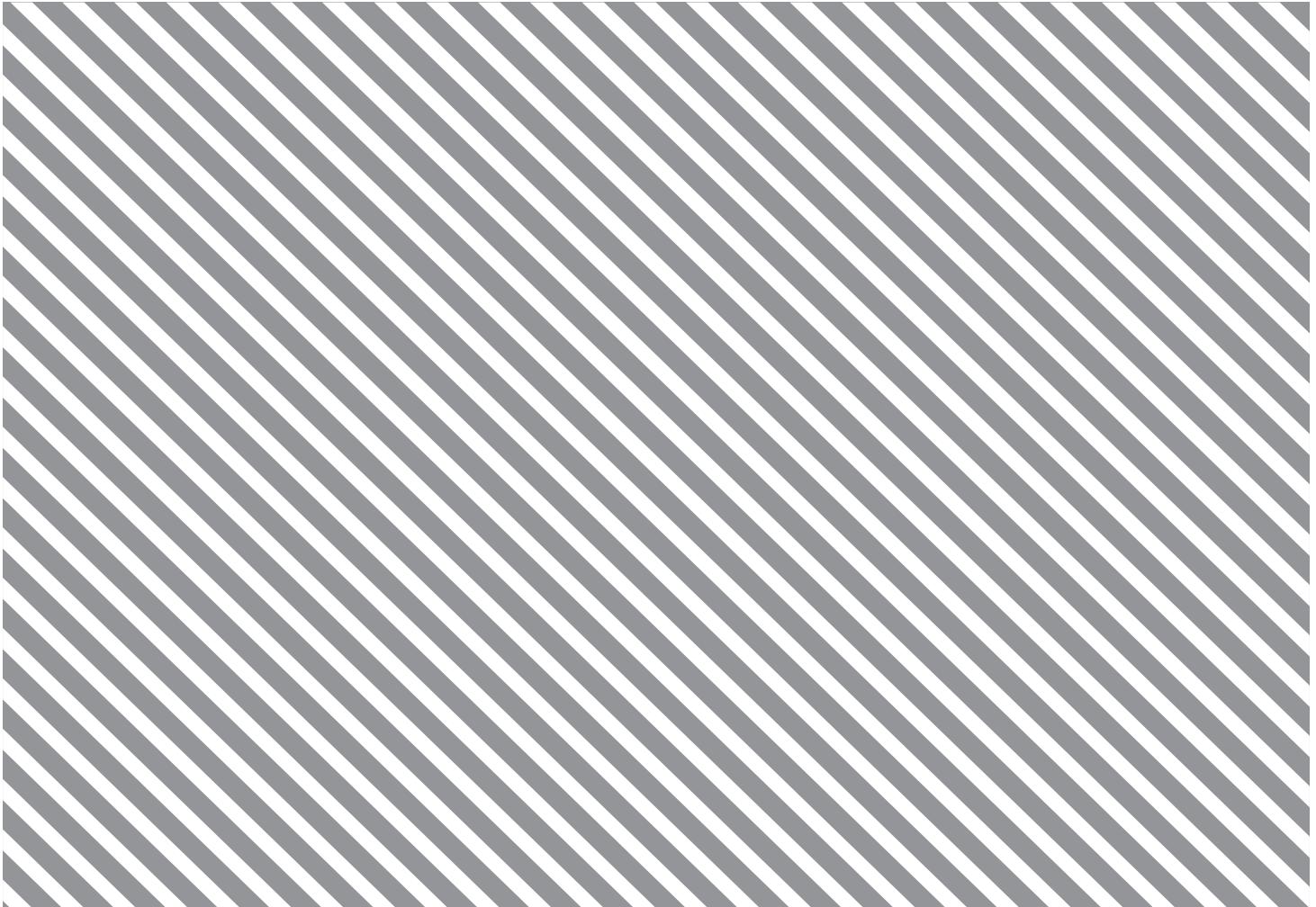


COMMITTED TO
IMPROVING THE STATE
OF THE WORLD

White Paper

Responsible Use of Technology

August 2019



World Economic Forum
91-93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland
Tel.: +41 (0)22 869 1212
Fax: +41 (0)22 786 2744
Email: contact@weforum.org
www.weforum.org

© 2019 World Economic Forum. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, including photocopying and recording, or by any information storage and retrieval system.

Contents

Foreword	5
Introduction	6
Section 1	7
Section 2: The entire value chain	9
Section 3: Act–Enable–Influence	17
Conclusion	18
Acknowledgments	19
Endnotes	20

Foreword



World Economic Forum, **Zvika Krieger**, Head of Technology Policy and Partnerships; Member of the Executive Committee

The Centre for the Fourth Industrial Revolution Global Network was launched early 2017 to co-create forward-looking policy frameworks and governance protocols, with broad multistakeholder engagement and buy-in, to accelerate the adoption of emerging technologies in the global public interest. The centre's work, in areas such as artificial intelligence and machine learning, blockchain, data policy, and internet of things (IoT), is driven by human-centricity, actionability, transparency and ethics.

Today, the pervasive use of technology has triggered a spirited debate on how new and disruptive technologies should be managed and governed. In particular, as government struggles to keep up with the unprecedented speed and scale of technological change, what is the responsibility of companies to self-regulate the technology they are designing and deploying?

This urgent and important challenge requires the engagement of stakeholders across the whole technology value chain, from the initial design and development to the sale of technology and its ultimate end use. Additional urgency is driven by the rise in employee activism on the ethical use of technology, as well as increased consumer awareness and concern that is beginning to affect the bottom line.

In our numerous conversations with c-suite leaders across the technology industry, we've learned that many companies lack even a basic framework for grappling with how their products are designed and who they should be sold to – and, in absence of a systematic approach, many of them are defaulting to reactive one-off decisions that are drawing increasing scrutiny.

To that end, the centre is developing a practical guide that companies and founders anywhere in the world can use to ensure that ethical considerations are integrated at key points in the product life cycle. It will include an evaluation schema, a catalogue of best practices, and a compilation of case studies – to be created through consultation with practitioners from business, government and civil society through a series of workshops across the Centre for the Fourth Industrial Revolution Global Network. The intent is to help corporate users address ethical issues associated with their technology and test which use cases might have adverse consequences.

While there have always been ethical choices to be made in business and governance, we now find ourselves in new territory. Internationally recognized laws and standards on human rights provide a common standard of achievement for all people in all countries, but more work is needed to understand how they can be best applied in the context of disruptive technology.

As a result, there is an opportunity for companies, national governments, civil society organizations and consumers to teach and learn from each other how to better build and deploy ethically sound technology. By doing so, we, together, are on the way to designing a better future.

This ambitious and inclusive vision requires the collaboration of all global stakeholders. The World Economic Forum, established 50 years ago with the mission to improve the state of the world, is the International Organization for Public-Private Collaboration. BSR (Business for Social Responsibility), a key partner on this project, is a global non-profit organization that works with its network of more than 250 member companies and other partners to build a just and sustainable world. Together, we are pleased to be contributors to building this new vision.

This paper draws upon the experience of a multistakeholder steering committee and is designed to propose solutions and stimulate further dialogue. We plan to refine the ideas presented in this paper over time and create accompanying toolkits and resources. We look forward to receiving your insights.



World Economic Forum, **Pablo Quintanilla**, Project Lead and World Economic Forum Fellow from Salesforce



BSR, **Hannah Darnton**, Program Manager of Ethics, Tech, and Human Rights



BSR, **Dunstan Allison-Hope**, Managing Director

Introduction

The increasingly pervasive use of technology in our everyday lives has triggered debate on how new and disruptive technologies – such as artificial intelligence (AI), robotics, 3D printing, internet of things (IoT), 5G, blockchain, quantum computing, autonomous vehicles, biotechnology and nanotechnology – should be managed and governed.

To contribute to the evolving dialogue and influence the management and governance of new and disruptive technology, this paper proposes a new framework for the responsible use of technologies. This framework is intended to connect the practical steps companies can take to address responsible use issues systematically with the underlying concepts of ethical thinking and international human rights standards.

We make the case that responsible use of disruptive technologies will benefit from both ethics-based and human-rights-based approaches, with active participation from technology companies, governments and other stakeholders across the entire technology industry value chain – from the earliest phases of creation to the application or use of technology by the end customer. Ethics and human rights approaches should not be thought of as oppositional but rather as two synergistic approaches for the responsible use of technology.¹

To do this, we lay out the three phases of the technology value chain – design and development; deployment and sale; and use and application – and discuss the necessary actions in each phase.

We aim to establish a common foundation for dialogue that allows technical and non-technical stakeholders to engage in the conversation, understand the questions that need to be asked and co-create approaches that institutionalize ethics and human rights into the development, deployment and end use of new and disruptive technology. We also propose key questions to ask at every stage of the technology value chain and a model indicating the main courses of action that companies can take to address responsible use issues.

We plan to use this paper and the dialogue it creates to inform the creation of implementation tools for organizations to advance responsible technology practices. This might include a responsible-use decision framework, resources to improve the integration of both ethical and

human-rights-based approaches across roles and business functions, and playbooks for each stage of the technology product life cycle.

Throughout this paper we refer to both risks and opportunities. We do this because we believe the power of disruptive technologies to address some of the world's most pressing environmental, social and governance challenges will benefit from the systematic consideration of opportunities and positive impacts, alongside the proactive prevention and mitigation of risks and adverse impacts.

Section 1

We believe the responsible use of technology will benefit from a smart and deliberate combination of both ethics-based and human-rights-based approaches. A human-rights-based approach provides a universal foundation upon which various ethical frameworks, choices and judgments can be applied.

Ethics-based approaches

- A framework for decision-making in situations where right and wrong, good and bad, are not clearly defined
- Address issues of fairness and social justice where different schools of thought and ethical standards exist; when various choices can be made, and different paths can be chosen
- Different traditions, cultures, countries and religions may choose different outcomes and priorities suited to specific needs and sensitivities

Human-rights-based approaches

- A focus on the experiences of the most vulnerable and a holistic recognition of what all members of society need in order to live with dignity and thrive
- Based on internationally recognized laws and standards; a common standard of achievement for all people
- Established rights that should always be protected and respected
- A minimum threshold and baseline expectation for the responsible use of disruptive technology
- An internationally endorsed framework for defining company responsibility that considers the critical role of governments

The ethics approach

The ethics-based approach takes a broad perspective on matters of right and wrong and includes the perspectives of many traditions, cultures and religions. Ethical terminology has been adopted by a wide range of stakeholders from government, business, academia and civil society in the context of many disruptive technologies with the potential for widespread societal impact, such as AI, robotics and biotechnology.

An ethics-based approach encourages a consideration of different views and is well-suited when public dialogue is needed to inform pathways forward.² Examples include different notions of fairness in access to public benefits, distributive justice and the ethics of when and how to go to war.³ An ethics-based approach allows for the identification of opportunities that are “socially acceptable or preferable, while at the same time potentially averting costly mistakes by elucidating what is social unacceptable”.⁴ Beyond the realm of the fundamental standards set by international human rights, ethical frameworks are especially useful when considering different perspectives across a range of geographies and jurisdictions, allowing for the recognition of different local contexts, priorities and cultural nuances.⁵

An ethics-based approach creates an entry point for conversation about the responsible development, deployment and use of technologies. The abstract nature of ethical principles allows for flexibility and innovation when analysing the potential risks and opportunities of disruptive technologies that go beyond regulatory compliance.

Some of the more widely noted ethics-based approaches to date have predominantly focused on AI. At least 25 countries have announced national AI strategies⁶, ethical task forces are forming around the globe, and institutes and associations ranging from the European Commission to the Association for Computer Machinery and the Institute for Electrical and Electronics Engineers have published ethical guidelines for AI or autonomous systems.⁷

Technology companies, including Google, Microsoft, Salesforce, IBM, NEC and Cisco, have also set out their ethical principles and positions, and are moving on to the important but challenging task of operationalizing these principles and integrating them into the business. They are beginning to address important ethics-based questions, such as who to sell to, what restrictions on product use to require of customers, what best practices and guidance to promote, and how to build different notions of fairness and distributive justice into algorithms.

Many fields of applied ethics and codified principles, institutions and cultures have emerged in response to disruptive technologies in the past. For example, in the 20th century, the domains of healthcare ethics, bioethics and research ethics emerged in response to new medical and biological technologies. Likewise, nuclear energy, technologies of mass destruction, aviation and space technologies all have ethical structures around them and strong formal bodies to institutionalize norms into policy frameworks and cultural expectations. Examples include treaties controlling nuclear energy and limiting nuclear weapons, institutions that coordinate aviation, and the Outer Space Treaty for governing space exploration.

The human-rights-based approach

One of the great achievements of the past 70 years has been the universal recognition that basic rights and fundamental freedoms are inherent to all human beings, inalienable and equally applicable to everyone, and that every one of us is born free and equal in dignity and rights. This has been accompanied by the creation of comprehensive international human rights law, setting out the obligations of governments to promote and protect human rights and establishing international mechanisms to protect these rights.

Furthermore, one of the great achievements of the past 10 years has been the development of the UN Guiding Principles on Business and Human Rights (UNGPs)⁸ as the authoritative global framework for business and human rights. The UNGPs clarify the responsibility of all businesses to respect human rights, which means to act with due diligence to prevent, mitigate and address human rights harm. The UNGPs also elaborate on the duty of governments to protect against human rights abuse from business activities and explain the right of those who have been harmed to an effective remedy. Endorsed by the UN Human Rights Council in 2011, the UNGPs build upon decades of human rights principles, standards and methodologies that are relevant for business.

A human-rights-based approach identifies rights holders (in this case, people who use or are affected by technologies) and duty bearers (in this case, companies or governments designing, deploying and using technologies) and considers the rights and responsibilities of each. A human-rights-based approach provides the mechanisms through which duty bearers can meet their obligations and ensure that rights holders have access to an effective remedy.⁹ A human-rights-based approach also requires attention to the needs and interests of vulnerable groups and marginalized populations.

In the case of disruptive technologies, international human rights standards and conventions provide a universally agreed rights – such as privacy, security, freedom of expression, non-discrimination, child rights and freedom of movement – against which companies (and governments) can assess the potential impact of disruptive technology in a structured, methodical and comprehensive manner.

By providing a common standard of achievement for all peoples, 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.¹⁰ In line with the UNGPs, a number of companies are undertaking human rights due diligence of disruptive technologies, such as AI, IoT and social media platforms, and providing new forms of access to remedy.

In recent years, several efforts have focused on the role of human rights in the design, deployment and use of disruptive technologies. For example, in 2018 the UN Special Rapporteur on the promotion and protection of the

right to freedom of opinion and expression shared his report on the human rights impact of AI systems with the UN General Assembly. Additionally, the UN Special Rapporteur on extreme poverty and human rights is preparing a report on the human rights impacts of digital technologies in the implementation of national social protection systems, to be presented to the General Assembly in October 2019.¹¹

National governments, including those of Australia, Canada and France, have also launched initiatives focused on the exploration of the human rights impacts of AI, while several others, including India and Mexico, aim to address affiliated issues, including social, economic and financial inclusion. The multistakeholder Global Network Initiative (GNI) was founded in 2008 to address the novel freedom of expression and privacy challenges arising from the global internet and to hold member companies (such as Google, Facebook and Microsoft) accountable for their human rights approach.¹²

Implementing the two approaches

We believe the responsible use of technology requires a smart and deliberate combination of ethics and human-rights-based approaches. In practical terms, this could mean, for example, that ethics and human rights reviews of disruptive technologies undertaken to inform company decision-making are combined into one process.

However, both the ethics-based and human-rights-based approaches have faced critique. For some, the ethics-based approach lacks a foundation in government and company accountability, and is viewed as the “easy” or “soft” option for companies.¹³ For others, the human-rights-based approach is limited in its ability to incorporate the very different notions of fairness, distributive justice, or social cohesion that can exist in different societies and local cultural contexts.

Considering the two approaches as complementary, rather than oppositional, will create more robust and holistic mechanisms to identify opportunities and mitigate risks. The two approaches are intended to do different things and, in combination, are more powerful for it. Human rights reinforce ethics, and ethics reinforce human rights – indeed, human-rights-based approaches draw upon many traditions of ethical thinking and represent universal principles that have been broadly endorsed across borders and cultures.

The human-rights-based approach is grounded in universally agreed international laws, norms and responsibilities of business; an ethics-based approach affords broader inclusion of issues such as fairness, distributive justice and cultural contexts. A human-rights-based approach provides a universal foundation upon which ethical frameworks, choices and judgments can be applied, while also offering a framework to address rights that may conflict with each other.

The Responsible Design, Development and Use of Technology project seeks to take a multistakeholder and deliberative approach to exploring and defining this complementarity in more detail.

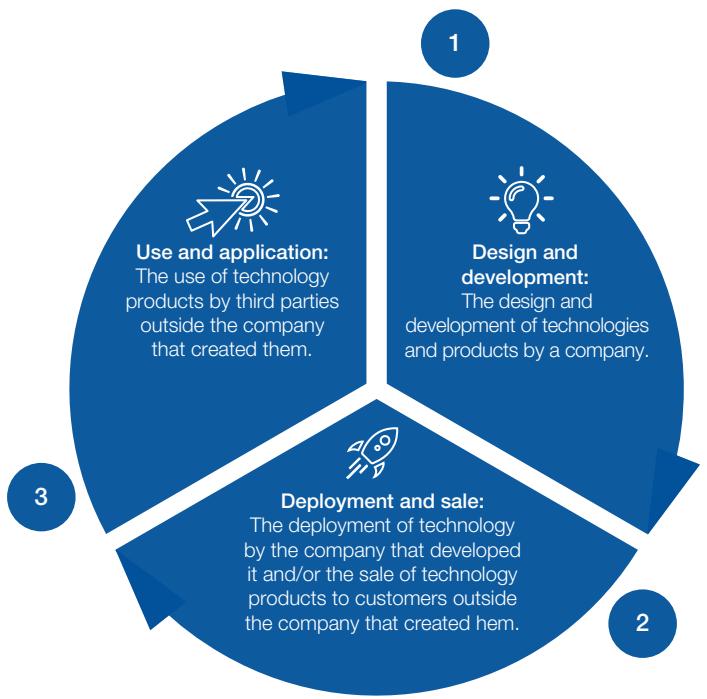
Section 2: The entire value chain

The life cycle of technology

Understanding the life cycle of technology will allow us to better appreciate the ethics and human rights opportunities and risks of the technology in question, as well as the responsibility, control and leverage held by different actors across the technology value chain.

There are three main stages in the life cycle of disruptive technology: how they are designed and developed; how they are deployed and who they are sold to; and how they are applied and used by an individual or entity other than the original creator.

Framing the life cycle of technology into these three distinct phases allows for the deliberate integration of combined ethics and human-rights-based approach at each step of the process, and the clear identification of the responsibility and leverage held by different actors and stakeholders throughout the life cycle.



The table below summarizes the distribution of responsibility by outlining the examples, key questions, actions and actors that can be considered across the three phases of the technology life cycle.

	Design and development	Deployment and sale	Use and application
Questions	<ul style="list-style-type: none"> – What is the socially beneficial purpose of this technology? – Should this technology exist? – What are the potential misuse cases, and can they be avoided, prevented, mitigated or remedied? – Who will this product affect? – Are there similar products to learn from? 	<ul style="list-style-type: none"> – Who should we sell or not sell this technology to? – Will the customer use the technology as intended? – Are there uses that should be disallowed or discouraged? – Can disallowed cases be prevented or mitigated? – Can the product be controlled after sale? 	<ul style="list-style-type: none"> – What are the social costs and benefits from using this technology? – What are the adverse human rights impacts arising, and how can they be prevented or mitigated? – What are the positive human rights impacts arising, and how can they be maximized? – What feedback can inform new design iterations?
Actions	<ul style="list-style-type: none"> – Ethics and human rights training – Human rights due diligence (e.g. via “human rights by design” methods) – Use tools for ethical thinking – Ethics review panels 	<ul style="list-style-type: none"> – Systematic human rights due diligence of potential use cases – Inclusive ethics dialogues 	<ul style="list-style-type: none"> – Human rights impact assessments of actual and specific use cases – Inclusive ethics dialogues
Prominent actors for each phase	<ul style="list-style-type: none"> – Universities and research institutions – Technology companies – Investors 	<ul style="list-style-type: none"> – Technology companies 	<ul style="list-style-type: none"> – End users – Consumers
Prominent actors within companies	<ul style="list-style-type: none"> – Board – Product and project managers – Engineering – R&D – Legal, public policy and ethics – CR and sustainability 	<ul style="list-style-type: none"> – Sales and marketing – Legal, public policy and ethics – CR and sustainability – Regulatory compliance – Export controls 	<ul style="list-style-type: none"> – Operations – CTOs, IT functions and systems integrators – Legal, public policy and ethics – CR and sustainability
Actors for every phase	<ul style="list-style-type: none"> – Civil society organizations – Company employees – Governments and regulators 		

Taking action

It is important at this point to understand the relationship between the notions of “responsibility” and “leverage” that apply to companies.

As set out in the UNGPs, the corporate responsibility to respect human rights means to avoid infringing on the human rights of others and to address any adverse human rights impacts with which a company is involved. Meeting this

responsibility requires taking *appropriate action*, including human rights due diligence to avoid causing or contributing to adverse human rights impacts through their own activities, and seeking to prevent or mitigate adverse human rights impacts that are directly linked to their operations, products or services by their business relationships.

Also set out in the UNGPs, leverage means the *ability a company has to affect change* in the wrongful practices of an entity that causes or contributes to a harm. Where

companies have leverage to prevent or mitigate harm they should exercise it; where they lack leverage, they should find ways to increase it, such as by collaborating with others.

Accompanying responsibility and leverage is the duty governments have to prevent, investigate, punish and redress human rights violations committed by companies through effective policies, legislation, regulations and adjudication – for example, by enacting and enforcing laws that require companies to respect human rights.

With these definitions as context, the following diagram sets out the course of action companies can take to prevent or mitigate adverse human rights impacts that the authors of this report have most commonly identified during real-world human rights and ethics due diligence engagements.

Companies that design, develop, deploy or sell disruptive technology – the vendors – have at least three main courses of action available to prevent or mitigate adverse human rights impacts. They can set and implement limits on what customers can or cannot do with disruptive technologies by establishing “acceptable use policies” covering impacts relevant for that technology, such as privacy, surveillance, or discrimination. They can define who they will or won’t sell to by creating whitelists (approved customers) or blacklists (blocked customers). And they can reduce the likelihood of product misuse by sharing guidance, training and best practices with others.

Companies that buy, use and apply disruptive technology – the customers – also have at least three main courses of action available to prevent or mitigate adverse human rights impacts. They can make proactive attempts to understand the real impacts arising from their use of disruptive technology by undertaking human rights impact assessments for their own use cases. They can make judgment calls and choices about how the disruptive technology will be used to avoid, prevent or mitigate impacts by acting upon the findings of the assessments. And they can be deliberate in communicating their lessons learned about product misuse and abuse back to the vendor.

There are also courses of action to prevent or mitigate adverse human rights impacts that exist across the entire value chain. Companies can engage in proactive transparency to increase collective awareness of how a technology works, with the aim of informing better decisions by others, such as users, governments and partners. Companies can advocate for standards, policies, laws and regulations from governments at all levels that define how technology should or should not be used. And companies can engage with a diverse range of stakeholders and deploy strategic foresight and futures methodologies (such as scenario planning) to anticipate adverse impacts that might otherwise go unnoticed.

Key Actions by Vendors

- **Acceptable Use Policies**
to define what customers can do and can't do.
- **White-Listing/Black Listing**
to define who you will and won't sell to.
- **Guidance and Training**
to share best practices and risk-migration priorities.

Key Actions by Customers

- **Understand Impact**
by undertaking human rights impact assessments.
- **Implementation Choices**
by deciding how the technology will be used.
- **Feedback Loops.**
that provide real-life insights back to the vendor.



Transparency

by sharing the science behind the technology and increasing shared understanding around how it works.

Standards

by collaborating with companies and stakeholders to set minimum standards for how technology should be used

Stakeholder Engagement

by engaging with rights holders (especially vulnerable populations and marginalized groups) throughout.

Regulation

by advocating for policies, laws and regulation for how technology can and cannot be used.

Strategic Foresight and Futures

by deploying human rights due diligence methods designed to uncover blind spots and anticipate uncertain futures.

It is noteworthy that these suggested measures have their shortcomings. A vendor may establish acceptable use policies – on data privacy, for example – but not have the insight necessary to enforce them effectively. There may be circumstances where society doesn't want companies to be deeply engaged with deciding who they do and don't sell products to, such as communications infrastructure that enables freedom of expression. And there may be situations when transparency heightens the risks faced by vulnerable groups. These shortcomings emphasize the importance of taking systemwide approaches to the responsible use of disruptive technology and not relying on the actions of a single company or government alone to affect change.

Design and development phase

The responsible design and development of technology requires a multidisciplinary approach in which a wide range of functions – engineering, product, policy, governance, sustainability and others – are all actively engaged. This often requires increased communication and collaboration between existing teams to build cross-functional knowledge.

While policy and sustainability teams are often involved in laying the foundation for the consideration and execution of ethics and human-rights-based approaches across an organization, engineering and product teams can be brought in to consult on the implementation of these processes, assist in initial assessments of products, services and technologies, and advise on the prevention and

mitigation of ethical and human rights implications through design modification.¹⁴ When companies might not have the expertise or know-how to fulfil their ethical and human rights obligations, they should bring in external experts and stakeholders.

The following four practices can assist companies in their efforts to further integrate ethical frameworks and human-rights-based approaches throughout design and development phase across a wide range of functions.

Awareness and recognition of ethical and human rights impact

The first step towards resolving any problem is to become aware of it and recognize that it exists and should be addressed. Too often industries engaged in disruptive technologies are slow to discover that they are having serious negative impacts on the world. Companies need to be aware of ethical and human rights impacts and, if they lack the ability to notice such impacts, they should hire employees or consultants who are. A human rights policy can establish the basis for raising human rights awareness among employees, supplemented by training with ethics and human rights experts. But these are only first steps; the responsible use of technology also requires the institutionalization of norms, tools and processes into corporate workflows and culture.

YouTube's recent ban on livestreamed broadcasts by children "unless they are clearly accompanied by an adult" is one example of an effort to prevent potential negative ethical or human rights implications. Noting the vulnerability of children in livestreamed videos across their platform, the Google-owned company has instituted new protection mechanisms, including AI-classifiers that find and remove content.¹⁵

Stakeholder and rightsholder engagement

It flows from the UNGPs that the responsible design and development of technology should include active consultation and participation of relevant rightsholders and stakeholders. Identifying the appropriate populations – especially vulnerable and marginalized groups who may be most impacted by the technology – and building engagement into the early phases of the life cycle of the technology allow companies to integrate findings and make appropriate adjustments to the design of the product to mitigate risks and maximize opportunities for positive impact.

Frameworks such as "value sensitive design", "human-centred design", "human rights by design"¹⁶ and "responsible innovation" provide roadmaps for how to identify relevant stakeholders and rightsholders throughout the design process, consider their interests and appropriately engage them to understand their values and the rights that may be implicated by the use of the technology.¹⁷ Microsoft has increased participation and inclusion of diverse perspectives through an "inclusive design methodology", including online toolkits "Inclusive

101" and "Inclusive Activities" to provide information on integrating inclusive design processes, stress-testing concepts, and engaging in long-term partnership to increase diversity of perspective.¹⁸

Fairness, accountability and transparency

Beyond direct engagement, several technical approaches have emerged in response to concerns about discrimination, bias and a lack of clear decision-making processes in AI. The use of diverse and representative data sets can mitigate bias and reduce discrimination in algorithms, while tools such as Accenture's fairness toolkit and IBM's AI Fairness 360 can help companies work towards fair outcomes and results. Companies can institute required minimum confidence thresholds prior to deployment to reduce inaccuracies.

Large, publicly available image datasets have similarly begun to provide options for more diverse and representative training materials, which can be useful for the development of machine-learning models. However, many of these images remain geographically skewed and incorrectly label images taken from under-represented regions. Google and the Conference on Neural Information Processing Systems Competition Track recently launched the Inclusive Images Competition on Kaggle to tackle this issue. By crowdsourcing images from around the world, Google aims to "geographically stress test" datasets and release a larger set of images for public use.¹⁹

In addition to representative datasets, diverse and representative development teams can provide alternate perspectives or new cultural contexts that help identify unanticipated issues or unintended use cases for review.

Airbnb has fought discrimination online by assembling a new product team of engineers, data scientists, researchers and designers specifically focused on reducing bias, while simultaneously bringing together external experts to outline a plan to "fight discrimination and build inclusion". As first steps in their redesign, Airbnb experimented with new layouts that remove photos of hosts, which previously appeared next to the images of a home, and increased the number of "instant bookings", which is believed to reduce the opportunity for discrimination based on a prospective guest's photo or name.²⁰

To increase inclusion and provide more relevant results for individuals of all skin tones, Pinterest also took steps to iterate on its platform. Working together, the Diversity and Inclusion, technical and engineering teams created and launched a new feature that allows users to filter results by skin tone.²¹

Adversarial testing or red-team testing

Adversarial testing is intended to push the product, service, or platform in question beyond current assumptions of how it will be used by "pressure testing" the use or application of each technology outside of its current geography, user base and easily foreseen use cases. While adversarial testing is frequently applied in the realm of AI and machine learning,

all technologies can be further tested prior to deployment and sale to prevent both intentional and unintentional discrimination bias, breaches in privacy and security, limitations on freedom of expression, physical harm and other adverse impacts.

Further application of adversarial testing can also provide an alternative to the “deploy and comply” strategy in which companies field-test systems and products on the public without sustained research on the effects, only addressing negative implications when noted by users or the general public.²²

An important consideration in this phase is whether a product, system or service “should be built at all, or when a prototype is too preliminary or unreliable to be unleashed on infrastructure such as hospitals or courtrooms”.²³ If both the likelihood and the severity of an ethical or human rights violation is high, the company will need to decide if that technology should still go to market.

In the past few years, we have seen a number of technology-enabled products released with vulnerabilities that put individuals’ health, security and privacy at risk. From a [pacemaker that could be hacked](#) to alter prescribed therapy for a patient, to a [sniper rifle that can be manipulated to change targets](#), and an [airplane model with faulty simulators](#), some products, should never have made it to the deployment and sale phase.

Available resources

Several initiatives exist to provide concrete action items and resources for engineers and designers throughout the development phase.

Technical organizations such as the Institute for Electrical and Electronics Engineers (IEEE) and the Association for Computer Machinery (ACM) have published ethical principles on autonomous systems, including the “Ethically Aligned Design” First Edition, to help in the translation from ethical principles to practices. It is noteworthy that the IEEE lists human rights as the first general principle.

Ethical OS, the Ethical Operating System, released in 2018, was built to help technologists, product managers and engineers prepare for the unexpected ways their technology may be used. The toolkit includes a checklist of eight areas of emerging risks and social harm, along with 14 scenarios to help teams think through long-term implications of their technology and future-proofing strategies to help take actions to mitigate or safeguard against misuse.

“Ethics in Technology Practice” is a training programme to teach technology ethics in corporate settings and consider how to integrate ethical thinking into corporate workflows. The Markkula Center for Applied Ethics at Santa Clara University also has a more general Framework for Ethical Decision-Making, as well as specific applied ethics resources such as cases, articles and tools for solving ethical problems in bioethics, business ethics, education

ethics, engineering ethics, environmental ethics, government ethics, internet ethics, journalism ethics, leadership ethics and technology ethics.²⁴

BSR has adapted its human rights due diligence approach to address the uncertainty inherent in disruptive technologies by adopting strategic foresight, scenario planning and other futures methodologies.²⁵

Deployment phase

In contrast with the design and development phase, the ethics and human rights considerations surrounding the deployment and sale of disruptive technology involve examining the customer, consumer and user relationships, and the potential ways a technology company can influence how their product, platform or service is utilized.

Sale to third-party company

When selling products, services or platforms to a third-party, companies have less control over the end use of the product but are still directly linked to ethics and human rights issues associated with the use of that product. Therefore, they should take steps to prevent and mitigate the risk of misuse, including by identifying and exercising any leverage they may have.

The extent to which risks of misuse can be mitigated will vary according to the company’s business model. For instance, a company selling “off the shelf” and generally available products will likely have less leverage than a company co-creating tailored solutions with customers or maintaining ongoing consulting relationships. Companies with subscription models, or those providing regular updates and maintenance to their platforms and products, may have some level of ongoing influence on usage.

There are at least three main prevention and mitigation strategies for technology companies: acceptable use policies; whitelisting or blacklisting customers; and providing training and guidance to customers. We describe each in turn below.

Acceptable use policies (AUPs)

A company can seek to influence how a product, service or platform is used by requiring the user to sign AUPs that establish the rules and guidelines on how a product, service or platform should or should not be used. Similar to terms of services, AUPs focus primarily on how the technology should be used and establish legal rules for the appropriate conduct of the user. It is important to note that while AUPs may provide “legal distance” between the technology company and the user by shifting liability to the user, it may not establish sufficient “moral distance” without further action as described below. Moreover, requiring a customer to sign an AUP is one step; enforcing the AUP, and taking action against violators, presents a range of challenges that can vary across technologies and business models.

Whitelisting and blacklisting

A company can seek to influence how a product, service or platform is used by restricting who it is sold to – either by operating “whitelisting” processes (where potential customers are required to achieve certain process milestones or minimum criteria to qualify as a customer) or “blacklisting” processes (where specific companies or entire sectors are eliminated as potential customers). A number of hardware companies – such as Ericsson, Nokia and Cisco – have well-established sales review processes to address risks of extreme surveillance, and internet companies – such as Microsoft, Salesforce and Google – are creating similar approaches to the sale of AI, big data and facial-recognition technologies.

The notion of restricting sales is not new. Technology companies already have sophisticated export controls that bar sales to restricted parties and flag transactions in which products and services may likely be misused.

ted measures into its internal compliance programme specifically to prevent goods exported from Europe being used for purposes associated with the violation of human rights, democratic principles or freedom of expression.²⁶ The European Commission has further proposed new regulations to invoke human rights considerations as a key justification for restricting exports.²⁷

While export controls are primarily about legal compliance, similar processes can usefully be deployed to go “beyond compliance” and fully integrate ethical and human rights considerations. Common features of these expanded processes include understanding (1) the nature of the customer, and (2) the nature of end use before making a sale.

The UK government’s 2014 report, *Assessing Cyber Security Export Risks*, provides detailed guidance on the incorporation of human rights considerations into export controls. Published in collaboration with techUK and the Cyber Growth Partnership, the report provides guidance and advice to cyber security companies working to “identify and manage the risks of exporting their products and services”.²⁸

Training and guidance

A company can seek to influence how a product, service or platform is used by providing guidelines and training on best practices’ Examples exist on everything from drones and automatic weapons to algorithmic decision-making. Embedding training and guidance into the sales process helps to ensure that end users have much improved knowledge, insight and capability to address ethics and human rights issues.

For example, in an effort to thwart the use of their ridesharing technology by “bad actors”, Uber provides its drivers with educational materials and training to identify signs of human trafficking.²⁹

Google has published recommended practices for AI³⁰ and an “inclusive ML guide” for customers using Google’s AutoML program to create their own customized machine-learning products.³¹ This includes guidelines on recommended fairness practices, information and example questions to assess fairness and bias, guidelines on training data and how to evaluate a model’s final performance. IBM’s AI Fairness 360, mentioned above, also provides a toolkit on bias metrics and mitigation techniques with industry-relevant policy specifications and tutorials.³²

DJI, the world’s largest drone manufacturer, following an alleged assassination attempt using one of its drones, has collaborated with governments and regulators to develop guidance and technological solutions to limit misuse. While this does not yet include explicit ethical or human rights considerations, it has taken a first step in training users on appropriate safety measures and exercising “good judgment”.³³

The Ethical OS and the Ethics in Technology Practice training programmes (also referenced above) offer customized training on the integration of ethical thinking throughout the life cycle of a technology or product and can create tailored risk checklists for specific products or teams seeking to mitigate improper or harmful use by third-party actors.

Professional associations, including the IEEE and ACM, continue to develop and promote training for their members. However, government, universities, civil society and investors can further amplify guidance and training by integrating them into regulations surrounding specific product use, incorporating them into university curriculums, and promoting cultural adjustments and expectations for how responsible behaviour can continue to evolve.

Limitations

While these strategies can result in positive shifts in the use of technology and can protect against ethical and human rights violations, their impact has two main limitations: the lack of certainty that they will work, as real control over product use resides with the customer; and the “substitutability problem” of a less scrupulous company making the sale anyway. The responsible company may do the right thing by not making the sale, but the impact on human rights will be neutral (or even negative) if another company steps in to fill the void. These limitations point to the need for more systemwide, legal and regulatory approaches that restrict uses of disruptive technology where they bring severe impacts, as well as the development of professional codes of conduct and expectations of behaviour.

Use/application phase

Once the technology company has done what it can during the development and deployment and sales phases, the primary (but not sole) responsibility resides with the company, government or consumer using the technology.

For example, in the transport industry, [airlines and airports are deploying facial recognition during the boarding](#) and screening process, raising issues on consent and non-discrimination. In the retail industry, [stores are deploying AI for theft protection](#), creating new privacy, security and discrimination risks, especially for vulnerable populations and marginalized groups. Energy companies are [deploying drones](#) to conduct regular maintenance checks, while food companies are collaborating [to use blockchain](#) to track food globally through supply chains.

There are an increasing number of consumer tools, including Consumer Reports' recently launched project Digital Lab, which will be developing new systems to test and report on the privacy and security of these digital products and services.³⁴ Consumers themselves are increasingly speaking out and exerting pressure on companies on these issues, and journalists continue to expose the potential harm of product misuse.

The implication of these trends is clear. While disruptive technologies are developed by the technology industry, they are being rapidly deployed across all industry sectors, including retail, financial services, agriculture, healthcare, transport and tourism, and it is essential that companies in these sectors review ethics-based questions and undertake human rights due diligence too. The responsible use of technology is not a problem for the technology industry alone to solve.

Section 3: Act–Enable–Influence

There are three major implications arising from the previous section: that systemwide approaches encompassing the entire value chain are needed; that company-only action will not sufficiently address ethics and human-rights-based challenges alone – policy, legal and regulatory approaches are needed too; and that companies will be unable to meet their responsibility to respect human rights if states neglect their duty to do so.

For this reason, companies can deploy a three-part approach based on BSR's "act, enable, influence" framework.

Act: Companies can "act" within their own company boundaries to create business strategies, governance and management approaches that directly address ethical and human rights standards and implementation practices. This is most relevant at the design and development stage.

Enable: Companies can "enable" integration of ethical and human-rights-based considerations beyond their own company boundaries by building positive, mutually beneficial relationships with external stakeholders and customers across the entire value chain. This is most relevant at the deployment and sale stage.

Influence: Companies can "influence" the adoption of responsible technology beyond their own company boundaries by promoting legal, policy and regulatory reforms that strengthen the relationship between commercial success and the achievement of ethics and human-rights-based outcomes. This can be done acting alone, or as part of multi-company and multistakeholder collaborative efforts that push for shared policies, solutions and outcomes.

By applying this three-part framework, companies can establish business strategies that embed the responsible use of technology into their core, create long-term value, and find ways to move beyond short-term performance pressures that often times prevents progress on ethical and human rights implications.³⁵

Responsible technology should not be the mandate of a single industry but a goal of society more broadly if we want to see a step change in both its creation and use. Collective action can pull together companies developing or engaging with technology within a specified industry. However, without a systemwide approach advocating for rights-protecting laws and regulations, increasing disclosure and transparency and providing best practice guidance for users, it will be difficult to realize a true shift in practices across industries and phases.³⁶

Here we illustrate the Act–Enable–Influence framework with the known case of climate change and the current case of facial recognition technologies.

Act–Enable–Influence applied to climate change

Act	Enable	Influence
Reducing a company's GHG emissions through energy efficiency or the use of renewable energy	Facilitating the reduction of GHG emissions by other companies by providing energy efficient technologies	Advocating for international/national agreements on climate change and government policies, laws, and regulations that result in reduced emissions

Act–Enable–Influence applied to facial recognition

Act	Enable	Influence
Eliminating bias from facial recognition systems and increasing accuracy	Only selling to reputable customers and providing training and guidance on best practice	Advocating for policies, laws and regulations for facial recognition that address human rights risk, such as how law enforcement agencies deploy facial recognition

Conclusion

Building on this report, the [Responsible Development, Deployment and Use of Technology](#) project seeks to produce a framework and suite of implementation tools for organizations to use to advance responsible technology practices. These tools will implement our desire for a smart and deliberate combination of both ethics-based and human-rights-based approaches. With a multistakeholder steering committee now in place, the project is focused on pursuing global stakeholder input and participation.

Planned deliverables include a “World Economic Forum Responsible Use Decision Framework” and a library of resources to improve the integration of both ethical and human-rights-based approaches across roles and business functions. Furthermore, the project will work to create toolkits or playbooks pertaining to each stage of the product life cycle: development, deployment and use.

Acknowledgments

This white paper has been published by the World Economic Forum as a contribution to a project, insight area or interaction. The findings, interpretations and conclusions expressed herein are a result of a collaborative process facilitated and endorsed by the World Economic Forum, but whose results do not necessarily represent the views of the World Economic Forum, nor the entirety of its Members, Partners or other stakeholders.

Lead authors

World Economic Forum, Pablo Quintanilla, Project Lead and World Economic Forum Fellow from Salesforce

BSR, Dunstan Allison-Hope, Managing Director

BSR, Hannah Darnton, Program Manager of Ethics, Tech, and Human Rights

World Economic Forum, Austin Hunter, Project Specialist

Contributors

Salesforce, Paula Goldman, Chief Ethical and Humane Use Officer

Salesforce, Christina Zhang, Program Specialist

Uber, Sabrina Ross, Global Head of Policy (Marketplace) at Uber

Deloitte Services LP, Jim Guszcza, Chief Data Scientist

IDEO, Kim Cullen, Interaction Design Director

Markkula Center for Applied Ethics at Santa Clara University, Brian Patrick Green, Director of Technology Ethics

UN Human Rights Office, Scott Campbell, Senior Human Rights Officer

Microsoft, Steve Sweetman, Director, Office of Responsible AI, Microsoft

Workday, Barbara Cosgrove, Vice President, Privacy Ethics Compliance; Chief Privacy Officer

Workday, Danielle Cass, Director of AI Ethics

Project Contributors

Omidyar Network, Yoav Schlesinger, Principal, Tech and Society Solutions Lab

Endnotes

1. "A Framework for Ethical Decision Making." Markkula Center for Applied Ethics at Santa Clara University, [www.scu.edu/ethics/ethics/resources/ethical decision making/a framework for ethical decision making/](http://www.scu.edu/ethics/ethics/resources/ethical_decision_making/a_framework_for_ethical_decision_making/).
2. This is not intended to imply that a human rights-based approach does not encourage a consideration of different views or public dialogue—on the contrary, these elements are essential to a human rights-based approach as well.
3. Allison-Hope, Dunstan, et al. "Don't Forget the Ethics: Blog." *BSR*, www.bsr.org/en/our-insights/blog-view/dont-forget-the-ethics.
4. "Governance with Teeth: How Human Rights Can Strengthen FAT and Ethics Initiatives on Artificial Intelligence." *ARTICLE 19*, www.article19.org/resources/governance-with-teeth-how-human-rights-can-strengthen-fat-and-ethics-initiatives-on-artificial-intelligence/.
5. "A Framework for Ethical Decision Making." Markkula Center for Applied Ethics at Santa Clara University, [www.scu.edu/ethics/ethics/resources/ethical decision making/a framework for ethical decision making/](http://www.scu.edu/ethics/ethics/resources/ethical_decision_making/a_framework_for_ethical_decision_making/).
6. Dutton, Tim. "An Overview of National AI Strategies - Politics AI." *Medium*, Politics AI, 25 July 2018, medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd.
7. "Ethically Aligned Design, First Edition: IEEE Standards Association." *IEEE*, ethicsinaction.ieee.org/.
8. "Guiding Principles on Business and Human Rights." *Ohchr.org*. Available at: https://www.ohchr.org/documents/publications/GuidingprinciplesBusinesshr_eN.pdf
9. Veen, Christiaan va, and Corinne Cath. "Artificial Intelligence: What's Human Rights Got to Do With It?" Data and Society: Points, May 2018. <https://points.datasociety.net/artificial-intelligence-whats-human-rights-got-to-do-with-it-4622ec1566d5>.
10. Allison-Hope, Dunstan. "Artificial Intelligence: A Rights-Based Blueprint for Business: Reports." *BSR*, www.bsr.org/en/our-insights/report-view/artificial-intelligence-a-rights-based-blueprint-for-business.
11. "Call for Input -Thematic Report to the UN General Assembly." *OHCHR*, www.ohchr.org/EN/Issues/Poverty/Pages/CallforinputGADigitalTechnology.aspx.
12. "GNI Home." *Global Network Initiative*, globalnetworkinitiative.org/.
13. "Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?" *Privacylab.at*.https://www.privacy-lab.at/wp-content/uploads/2018/07/Ben_Wagner_Ethics-as-an-Escape-from-Regulation_2018_BW9.pdf.
14. "Assessing Cyber Security export risks." [Https://Www.techuk.org/Images/CGP_Docs/Assessing_Cyber_Security_Export_Risks_website_FINAL_3.Pdf](https://Www.techuk.org/Images/CGP_Docs/Assessing_Cyber_Security_Export_Risks_website_FINAL_3.Pdf).
15. "Spangler, Todd, and Todd Spangler. "YouTube Bans Kids From Live-Streaming Video Unless Accompanied by an Adult." *Variety*, 8 July 2019, variety.com/2019/digital/news/youtube-bans-kids-live-streaming-1203231457/.
16. Allison-Hope, Dunstan, et al. "Human Rights by Design: Blog." *BSR*, www.bsr.org/en/our-insights/blog-view/human-rights-by-design.
17. "There Is a Blind Spot in AI Research." *Nature News*, Nature Publishing Group, www.nature.com/news/there-is-a-blind-spot-in-ai-research-1.20805.
18. "Microsoft Design." *Microsoft*, www.microsoft.com/design/inclusive/.
19. "Introducing the Inclusive Images Competition." *Google AI Blog*, 6 Sept. 2018, ai.googleblog.com/2018/09/introducing-inclusive-images-competition.html.

20. McGee, Chantel. "How Airbnb's Redesign Aims to Combat Discrimination on the Service." CNBC, CNBC, 10 Apr. 2017, www.cnbc.com/2017/04/07/airbnb-experimenting-with-site-design-to-fight-discrimination.html.
21. Lee, Nicole. "Pinterest's Diverse Workforce Helped It Design a Better Skin Tone Filter." Engadget, 24 Jan. 2019, www.engadget.com/2019/01/24/pinterest-skin-tone-search-diversity/?guccounter=1&guce_referrer=aHR0cHM6Ly93d-3cuZ29vZ2xLmNvbS8&guce_referrer_sig=AQAAAKYoA4W1NjsJdAMgWEQeogcW2KONFh6Nd6OpUiEIAnkTDqleK-c7oyPk1qBvf2WR24gQnZ21AcGuldJkfC7a_8pRJ6NuR5XCke49PWDpnJsY-5QIZj6rZnUoSUVlr8D8Cqj6yjdf2adMtetU-vollkxR5mGQZ9mbwlCFHwnlE2Qrye.
22. "There Is a Blind Spot in AI Research." Nature News, Nature Publishing Group, www.nature.com/news/there-is-a-blind-spot-in-ai-research-1.20805.
23. "Governance with Teeth: How Human Rights Can Strengthen FAT and Ethics Initiatives on Artificial Intelligence." ARTICLE 19, www.article19.org/resources/governance-with-teeth-how-human-rights-can-strengthen-fat-and-ethics-initiatives-on-artificial-intelligence/.
24. "A Framework for Ethical Decision Making." Markkula Center for Applied Ethics at Santa Clara University, [www.scu.edu/ethics/ethics resources/ethical decision making/a framework for ethical decision making/](http://www.scu.edu/ethics/ethics/resources/ethical_decision_making/a_framework_for_ethical_decision_making/).
25. Allison-Hope, Dunstan, et al. "Are You Prepared for the Future of Business and Human Rights?: Blog." BSR, www.bsr.org/en/our-insights/blog-view/are-you-prepared-for-the-future-of-business-and-human-rights.
26. Government of the Netherlands, Central Office for Import and Export. "Internal Compliance Programme." Internal Compliance Programme, vol. 1, 2018. <https://www.government.nl/documents/directives/2018/06/04/guidelines-for-compiling-an-internal-compliance-programme>
27. Machiko KANETAKE, Machiko. *The EU's Export Control of Cyber Surveillance Technology: Human Rights Approaches*. www.cambridge.org/core/services/aop-cambridge-core/content/view/511F6A505411E6BEBD4137C67CC68F7D/S2057019818000184a.pdf/eus_export_control_of_cyber_surveillance_technology_human_rights_approaches.pdf.
28. "Assessing Cyber Security export risks." [Https://Www.techuk.org/Images/CGP_Docs/Assessing_Cyber_Security_Export_Risks_website_FINAL_3.Pdf](https://www.techuk.org/Images/CGP_Docs/Assessing_Cyber_Security_Export_Risks_website_FINAL_3.Pdf).
29. "Uber's Commitment to Disrupt Human Trafficking: Uber Newsroom US." Uber Newsroom, 30 Jan. 2018, www.uber.com/newsroom/ubers-commitment-disrupt-human-trafficking/.
30. "Responsible AI Practices." Google AI, ai.google/responsibilities/responsible-ai-practices/.
31. "Inclusive ML | Google Cloud." Google, Google, cloud.google.com/inclusive-ml/.
32. "AI Fairness 360." IBM Developer, developer.ibm.com/open/projects/ai-fairness-360/.
33. "Fly Safe – Drone Flying Tips, Policies & Regulations, and More – DJI." DJI Official, www.dji.com/flysafe.
34. John, Allen St. "Consumer Reports Launches New Digital Lab." Consumer Reports, www.consumerreports.org/privacy/consumer-reports-launches-digital-lab/.
35. Redefining Sustainable Business. www.bsr.org/reports/BSR_Redefining_Sustainable_Business.pdf.
36. Redefining Sustainable Business. www.bsr.org/reports/BSR_Redefining_Sustainable_Business.pdf.



COMMITTED TO
IMPROVING THE STATE
OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744

contact@weforum.org
www.weforum.org