WORLD
ECONOMIC
FORUM

The AI Governance Journey:
Development and Opportunities

INSIGHT REPORT

OCTOBER 2021

Contents

About the Global AI Council

Since May 2019, the World Economic Forum Centre for the Fourth Industrial Revolution has been convening an informal group of government, business, academic and civil society leaders to help shape international governance and cooperation on artificial intelligence (AI) in the global public interest. At a time of rising public interest in the technology's social and economic implications, the Council set out to shape the public policy and corporate governance innovation agenda. Members meet virtually twice a year to discuss how to build trust and deepen understanding about the most appropriate and effective forms of AI governance and cooperation.

The Council is co-chaired by Brad Smith, President of Microsoft, and Kai Fu Lee, Chairman and CEO, Sinovation Ventures. Its membership comprises ministers and heads of regulatory agencies, chief executives, leading technical and civil society experts from the countries in the network and beyond.

The Council serves as an integral partner of the Centre's AI and Machine Learning Platform, helping to provide strategic guidance, shape its direction and identify governance gaps that its multistakeholder approach is particularly well-positioned to address.

Foreword

Brad Smith
President, Microsoft
Co-Chair, Global AI Council

The rapid advance in the uses of AI is remarkable. We have seen AI used to fight COVID-19, accelerate progress in environmental sustainability efforts and drive innovation in business. However, the technology is also challenging us with new and complex ethical issues and racing ahead of our ability to govern it. This is problematic because for all of us to benefit from AI, it needs to be built and used in ways that warrant people's trust.

Governments and companies around the world have been busy over the past few years developing AI principles. But what they are finding now is that the really hard work begins when they try and turn those principles into practices. Principles are important, but the only way to adhere to them is by developing the sound practices, tools and systems that make it easy for the people who have to build and use the technology.

This is why the work of the Global AI Council at the World Economic Forum has been so important. We have been able to bring representatives from the public, private and non-profit sectors to sit down together and have the larger societal conversations needed to help policy-makers, decision-makers and the general public understand the risks we face and how to ensure that our ability to govern the technology stays apace with the technology itself.

It has been an honour to serve as co-chair of the Council along with Kai-Fu Lee. We will continue to see new challenges arise from the advancement of AI, but the recommendations in this report, and the work we will continue to do, will help us collectively ensure that AI benefits all of society.

Kai-Fu Lee
Chairman and CEO,
Sinovation Ventures,
Co-Chair, Global AI Council

By most measures, the past few years have been significant for the advancement in technology, both in terms of breakthroughs and the speed of adoption.

From Deepmind to Open AI's GPT-3, these developments are tremendous, allowing companies and universities to replicate technology in months. GPT-3, especially, and the idea of pre-training data are developments that could be applied to genomics and sequencing. They represent the largest advancements in deep learning since deep learning itself. The challenge is that not everyone has access to the same types of resourcing.

AI's potential to do good appears to go beyond just the realms of climate and health; it is increasingly showing strong potential in agriculture and education. Hunger can be mitigated with this technology. We thought robots would be the next big thing, but it is hard to replicate the dexterity of human fingers and as a result, it is difficult to make a generalized product. Using AI for teaching is also advancing. Chinese ed-tech is taking off with game-like AI to target children's weakness and customize their education in ways that make it fun for kids; routine coursework is turned into a game. More and more teachers are not teaching AI but using AI for teaching, using interactive, gamified training to help develop in a personal way.
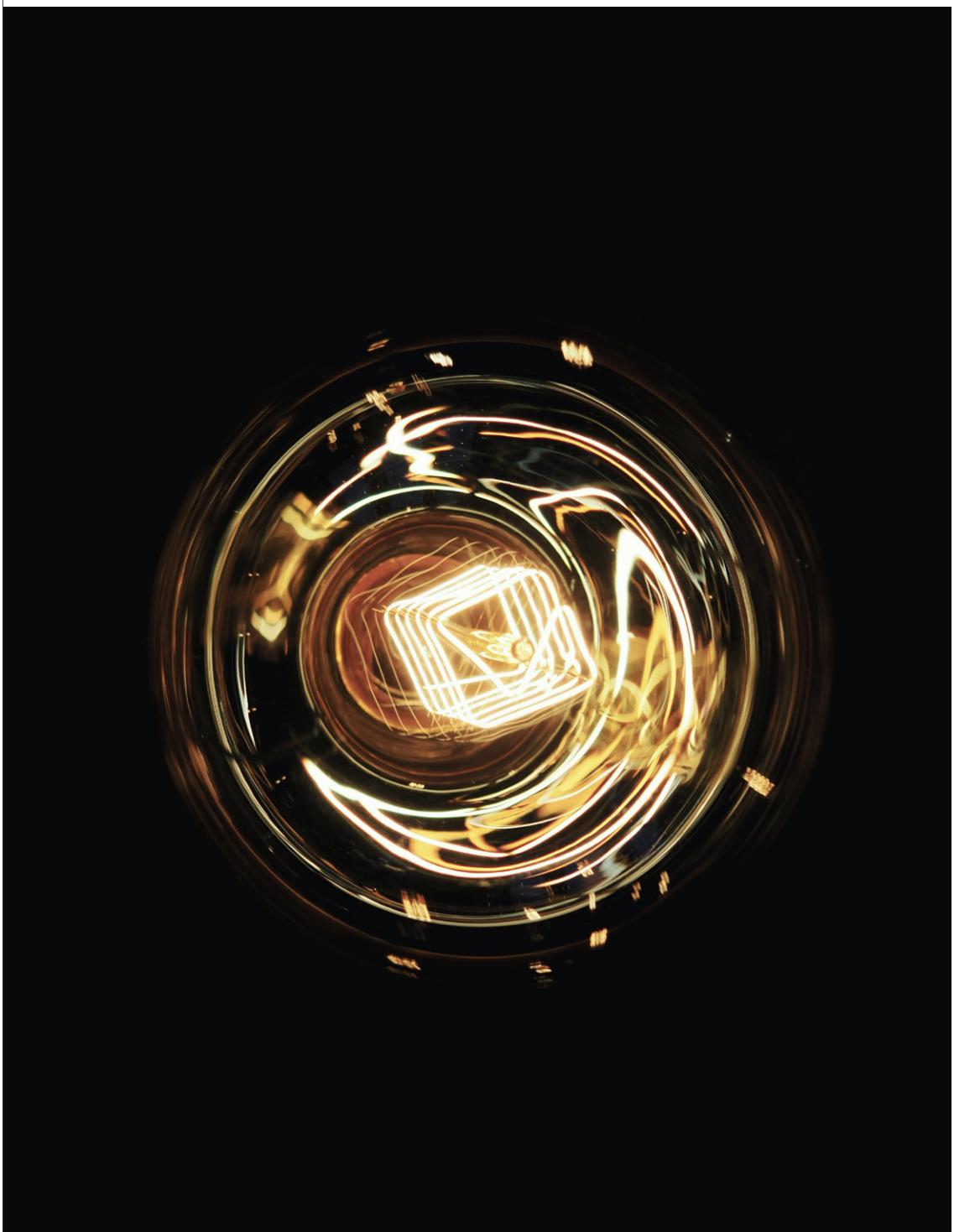
As co-chair of the Global AI Council at the World Economic Forum, I have enjoyed working with a strong community of leaders who are equally committed to ensuring that AI lives up to its promise to help humanity in so many ways. At the core of that commitment is making sure we get the governance right to help ensure the transformation is as smooth as possible. I am looking forward to seeing what more we can accomplish in the year ahead.

1 Introduction

The field of AI governance has grown rapidly over the last five years; the immediate challenge is to ensure responsible application and practices.

As artificial intelligence (AI) continues to appear in all aspects of our daily lives, the work of responsibly governing its use is more important than ever, to ensure that AI ends up doing more good than harm. The journey to construct this ecosystem of oversight has accelerated considerably in the last five years as leaders globally, across jurisdictions and sectors, have acted to ensure the world harnesses the full benefits of AI.
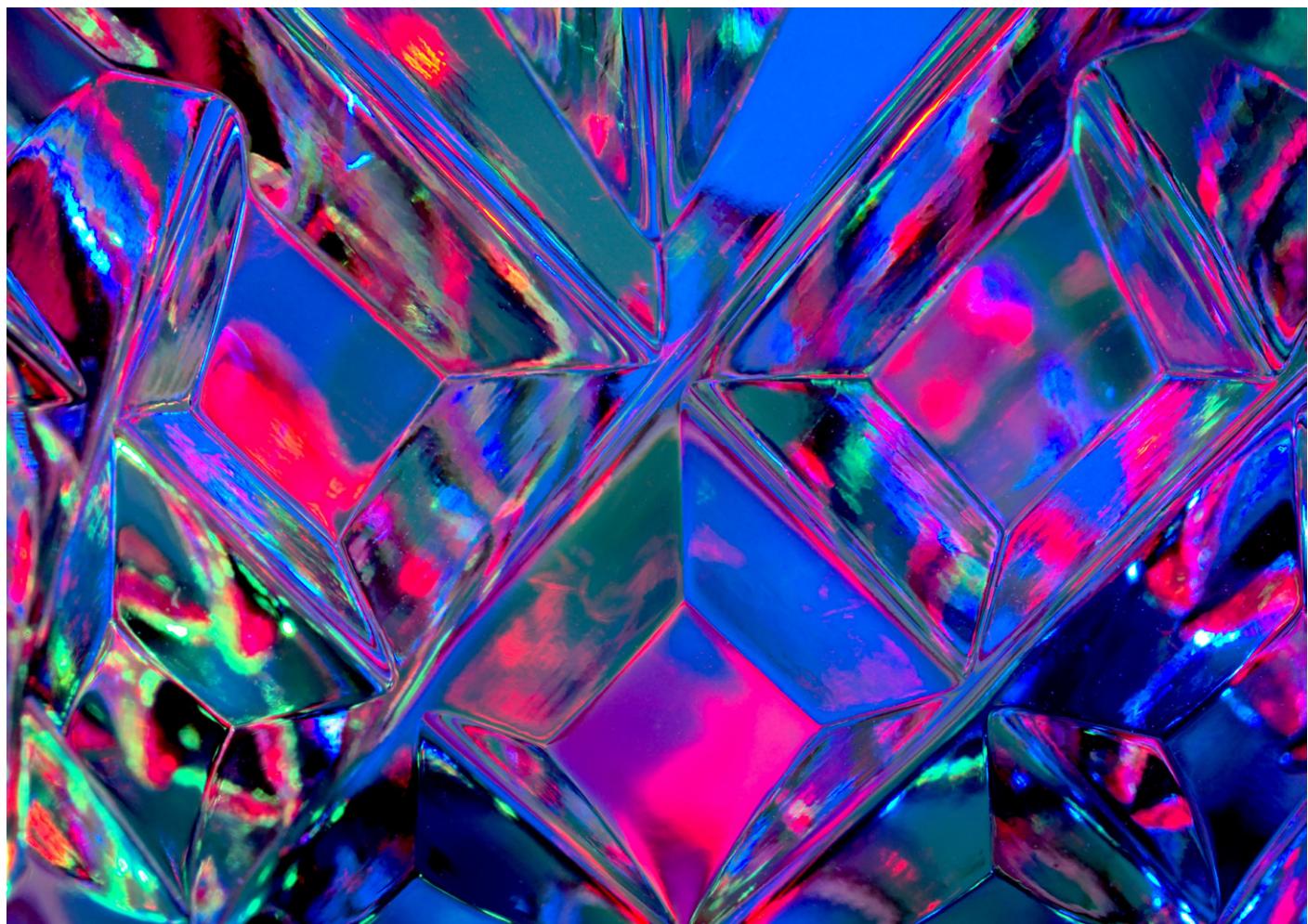
Harnessing those benefits requires prudent governance just as it did for previous technological advancements. But AI presents its own unique risks that are not easily dealt with using readily available policy tools. Competitive pressures threaten to create a collective action problem as private actors try to out-do one another even if it leads to suboptimal outcomes. At the same time, differing global perspectives on values such as privacy and fairness raise the potential of balkanization, curbing AI's potential.

In response, use of AI has required society to adapt and broaden its approach to safeguarding itself from the potential risks associated with advanced decision-making systems. One of these approaches has required the fostering of robust international and intersectoral cooperation. After the drafting and adoption of hundreds of high-level ethical AI principles, we began to see an emerging consensus around critical issues of importance. Eventually, a body of work emerged that offers useful strategies

to operationalize such guiding ethical beliefs. The explosion of governance frameworks over the past two years has been crucial in helping leaders from government, the private sector, civil society and the public to better understand the issues surrounding AI, including potential for fairness and discrimination, disparate impact, and associated issues of transparency and accountability.
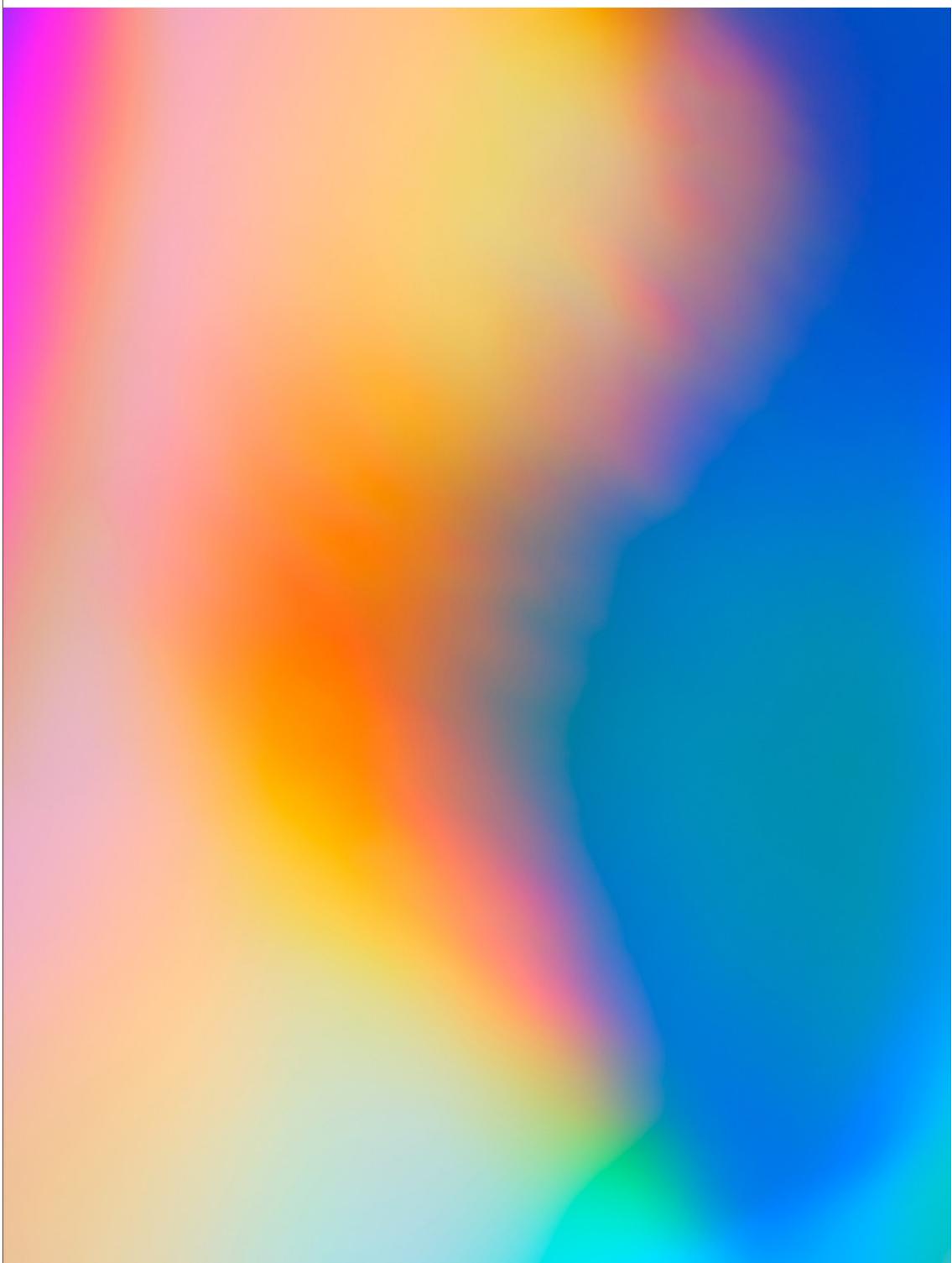
Despite this progress, much more innovation is required in the realm of AI governance if we are to both keep pace with the advancing capabilities of AI-based systems and orient its progress in a positive direction. Now more than ever, we need to apply a multistakeholder approach combined with an agile methodology to the design of AI governance. Efforts to responsibly govern the application of AI to ensure it benefits and protects all in society are in a sprint but will ultimately fall short if we do not act now.

There is tremendous opportunity to build upon and scale innovative approaches to AI governance at this critical juncture by sharing global best practices for responsible AI. This insight report aims to provide governments, business executives and other stakeholders a clearer picture of the emerging landscape and why their participation matters. With input from members of the World Economic Forum Global AI Council, it aims to summarize these contributions, highlight best practices and provide recommendations for moving forward.

(2) # Guidance on AI Governance from Global Council Members

" When we talk about the responsible governance and oversight of AI, really we are talking about strategy, new opportunity, market positioning and competitiveness; this goes far beyond questions of compliance and risk management. A compliance lens is critical, of course, but if it is the only lens, it is a trap. For AI to be successful out in the world, making the most of responsible use also requires a strategic lens and mindset.

Karen Silverman, CEO, Cantellus Group

" Trust is going to be the key currency of all companies in an increasingly digitized and AI- driven world. Agile AI governance should not simply be seen as a set of guardrails, but rather as a way to create more trust in a company among all its stakeholders – customers, employees, partners, regulators and civil society. Smart companies, like Germany's Deutsche Telekom, are realizing that trust in the use of their AI and technology can be a strategic competitive advantage today.

Simon Greenman, Co-Founder and Partner, Best Practice AI

" AI deployed in the real world must be effective, safe and deliver equitable results for all users. Teams building AI need to consider fairness and should strive to reduce risks of bias at all points of the development cycle. This starts with ensuring they provision for high-quality AI training data that is optimized to minimize bias by providing responsibly sourced data from a diverse and skilled global group of people.

Mark Brayan, CEO, Appen

" Society's reliance on AI goes hand-in-hand with external pressures to improve cybersecurity. Machine learning models need large data sets, which must be protected from unauthorized access and manipulation if we expect these solutions to maintain trust and confidence in the public's eye. Ensuring good cybersecurity practices must be at the top of the list for any executive or governing body working with AI.

Santeri Kangas, CTO, CUJO AI

" The stochastic nature of current data-driven approaches to AI is at the basis of most of the risks – in opposition to rule-, or model-based approaches, which are more amenable to formal verification and thus to trust. Most of the concerns with trust and risk in AI originate from the uptake of data-driven methods.

Virginia Dignum, Professor, Ethical and Social Artificial Intelligence, Umeå University

" We are living in a world where 'experience' is counted as the new currency. One of the fore-bearers of this experience comes from data insights. Data, which is now termed as the new "fuel" coupled with Insights, from "AI", which is now termed as the new "Electricity". AI, however, has its boundaries and these boundaries would become profound because of the vast data, variations and the silicon base. We live in the world where we are staring at the demise of the transistor. The current global dependency on energy for technology has become unsustainable so a fundamental shift in the mechanism of computing is needed. The technological world has a unique opportunity to combine the power of AI and Quantum in a responsible way to herald a new future, a future which would have machines wedded to human emotions and working in tandem.

C.P. Gurnani, CEO and Managing Director, Tech Mahindra

" How can we enlist the iterative nature of AI in aid of continuous improvement, bias reduction, consistency of alignment between desired and ethical outcomes and actual results? This is an area of emerging import.

Lenny Stein, Senior Strategic Advisor (2020-2021), Splunk

(3) AI Governance Eras

The journey to building trust in AI systems

With each industrial revolution come both opportunities and challenges, forcing society to reexamine whether current governance structures effectively offer a balance of sufficient oversight and the space for innovation. The Fourth Industrial Revolution is no exception. New risks require new safeguards.

As the engine that powers the Fourth Industrial Revolution, AI and the efforts to give it some guardrails sit at the core of many of these challenges. Reflecting on this journey – from when the term "artificial intelligence" was first coined in 1956 to its present renaissance – is important as we examine where we are and the work ahead.

3.1 | Pre-2010: AI winter and thaw

In its early years, it was expected that should AI develop and migrate into positions of power and responsibility, it would require extensive and thoughtful consideration over its governance. This stemmed from a growing awareness of the challenges posed by the exponential nature and pace of technology advancement. Even as AI endured what many historians have called AI winters, periods in which both progress and public interest in AI waned, there was always a mixture of enthusiasm and anxiety for what it would bring as well as uncertainty over whether present governance structures were up to the task. The extremely wide range of potential AI applications also meant that its societal implications would be just as extensive, raising concerns over privacy, security, bias, fairness, increased inequality, among others. Approaches, therefore, would need to draw on existing frameworks across many domains, including international law and public policy at a macro-level, and data governance, intellectual property and ethics codes that guide most computing and data scientists as well as engineers at a micro-level.[1]

3.2 | 2010-2016: AI acceleration presents challenges

Several breakthroughs in AI began to pick up after 2010 due to a combination of advancements running in parallel: higher levels of computing power, increased access to cloud computing and an explosion of digital data. Progress in deep learning approaches made it especially effective.[2] As in decades past, these advancements generated enthusiasm around the potential for AI-based systems to perform duties that would increase productivity, improve social services and strengthen our ability to tackle major global issues such as climate change and disease migration. Its potential to transform nearly every aspect of our lives from how we communicate, learn and transact business promised to completely revamp the digital economy as well as the physical one.[3]

Though exciting, the speed of AI development and deployment into the real world also generated a host of concerns, including the increased use of data and what that meant for privacy. In the rush to acquire or develop deep-learning technology, we saw some technology companies establish ethics boards to prevent abuse.[4] Some of the most immediate concerns arose with regard to its decision-making or decision-influencing applications, especially in sensitive areas such as law enforcement, human resources and finance. A few key studies and news stories surfaced about how AI-based systems can unfairly target certain social groups.[5] But macro-level concerns also arose, including severe labour market disruptions stemming from task encroachment and increased geopolitical rivalries.

Addressing these issues in a way that balances safety and innovation became a major priority for all those hoping to accelerate the benefits of AI. As the decade advanced, and as AI continued to accelerate, these governance gaps came to be understood as yet another symptom of the "Pacing Problem" – the decades-old idea that laws and regulations may not be capable of keeping up with the current pace of technology. As a result, we must broaden our conception of the governance structures needed to mitigate the risks posed by technology.

3.3 | 2016-2019: Principles and guidelines

❝ The initial response from industry and government was to develop or adopt sets of principles to guide AI development.

By 2016, there was heightened awareness that without proper safeguards to help guide its development, AI could disproportionately disrupt livelihoods, amplify social inequities, entrench existing biases or result in a reduction in privacy, all of which would weaken the public's trust in AI and quickly derail its potential to benefit society. The Executive Office of US President Barack Obama (2009-2017) published two reports, both raising concerns over many of these issues and making the case for why AI should be an urgent priority.[67]

The initial response from industry and government was to develop or adopt sets of principles to guide AI development. One such set of principles developed at the Asilomar Conference on Beneficial AI (USA) in January 2017 helped bring awareness to the various concepts and ideas that government and business had been grappling with until then, including privacy, fairness, non-discrimination, transparency, safety and accountability. This lit a spark around the world as organizations began publishing their own sets of principles around AI adoption. We also saw the founding of a few organizations and initiatives oriented around building AI that benefits all of society, including the OpenAI Institute in 2015 and the Partnership on AI (PAI) in 2016. The Ethics and Governance of AI Initiative and the World Economic Forum Centre for the Fourth Industrial Revolution were both founded in 2017.
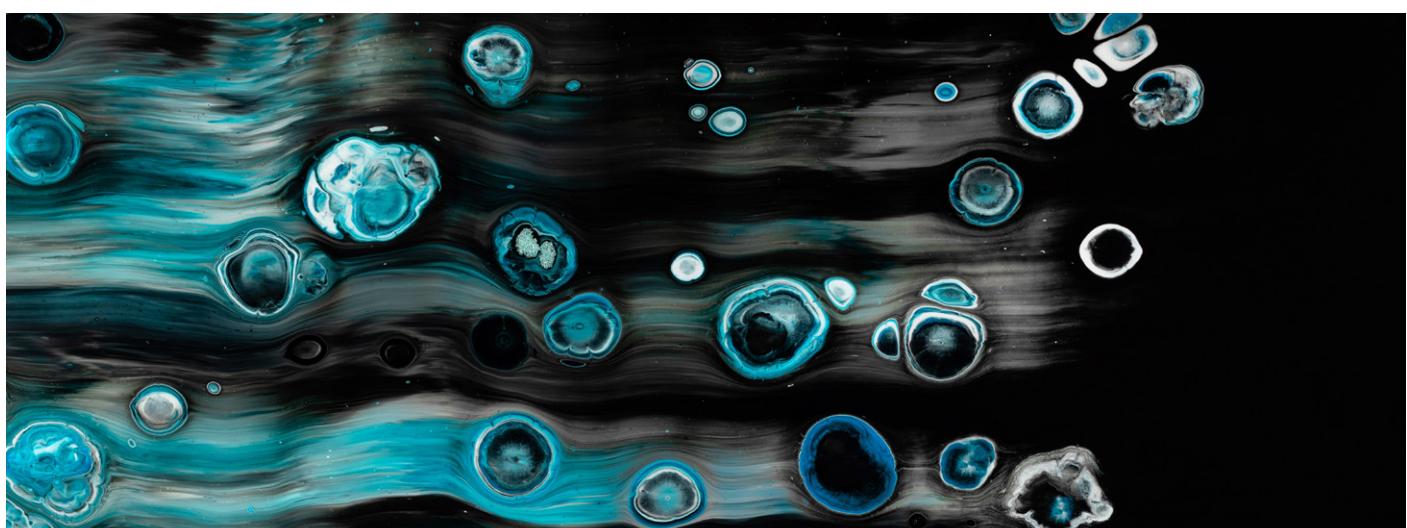
3.4 | 2019-present: Continued acceleration and governance innovations

The proliferation of AI principles and ethics guidelines helped reveal what a transnational approach to AI governance might entail. One analysis of major ethical guidelines found that aspects of accountability, privacy or fairness appear in about 80% of all guidelines.[8] Another analysis found eight areas of common concern: privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility, and promotion of human values.[9] Following the emergence of AI ethics principles, 2019 served as an inflection point that succeeded in both highlighting implementation gaps and orienting a significant number of efforts around translating principles into operational and actionable governance mechanisms on the level of society, individual organizations and down to development teams. Much of these efforts focused on "soft" governance mechanisms that focus on shaping norms around the use and adoption of AI, even though there was movement in more traditional forms of governance, ranging from outright bans on certain applications to proposed legislation that encourage algorithmic accountability.
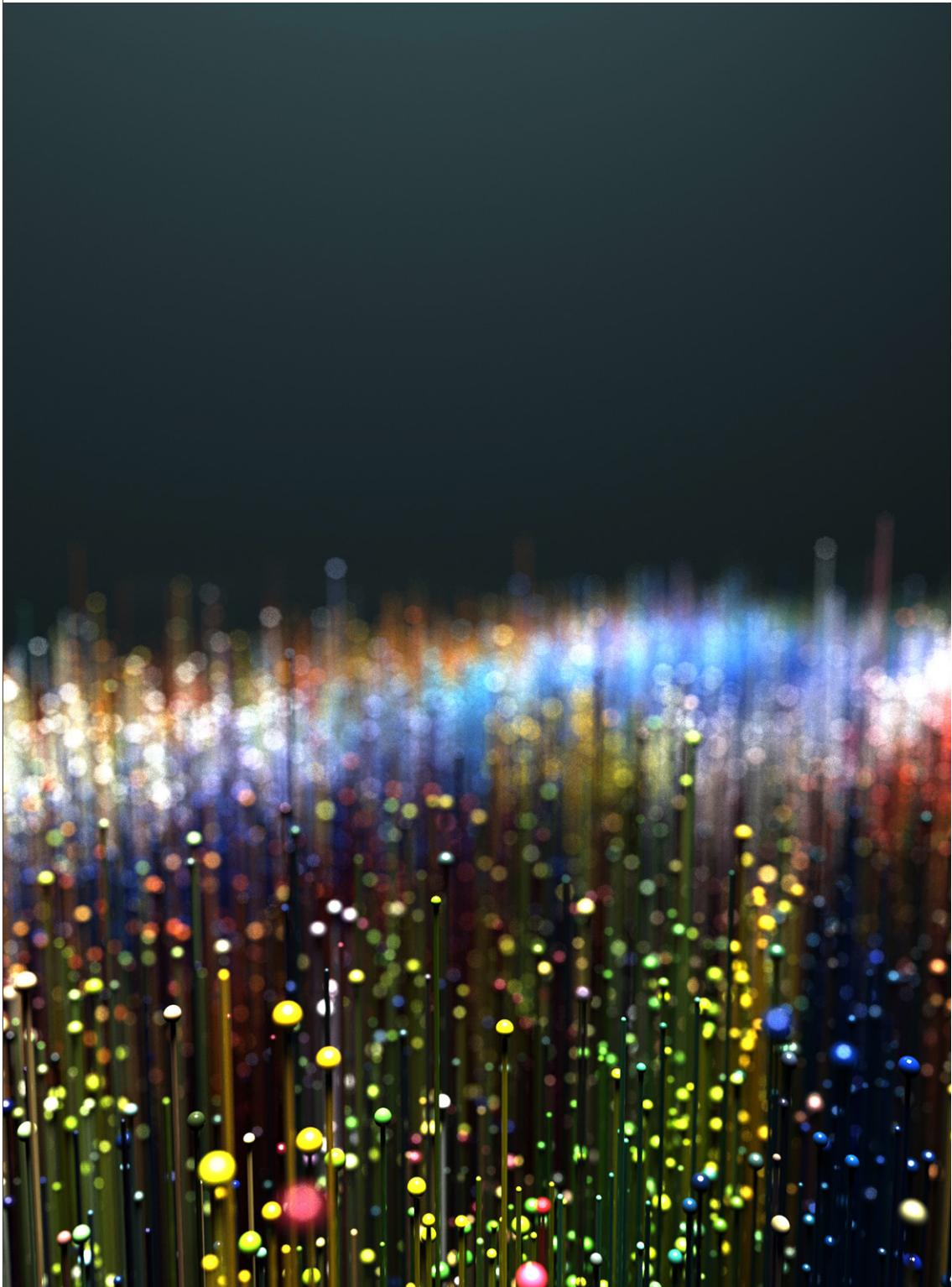
The attention devoted to this issue by such a wide array of players shows there is recognition that existing regulations and norms are insufficient to ensure responsible AI development. Moreover, the trend of exponential growth is likely to continue with recent research from OpenAI on AI and compute[10] and algorithmic efficiency[11] showing a reduction in the compute needed to train a specific capability. In other words, the pacing problem is as important as ever, if not more urgent. Advancing AI governance must therefore become a critical priority area for public- and private-sector leaders, as well as all those aiming to ensure that AI advances responsibly.

(4) From principles to practice

Operationalizing AI ethics to manage the risks without stifling innovation

By 2020, well over 100 ethical guidelines had been published by government actors, advocacy groups, international organizations, private companies and others.[12] The work of producing these represents an impressive area of progress in international cooperation on AI. The pressure to put these principles into practice, however, grew considerably as many companies soon struggled to navigate around mounting "techlash" and the uncertainty of how AI should be governed. In Deloitte's recent *State of AI in the Enterprise* survey, a majority of global respondents said their organization is slowing adoption of AI technologies because of the risks. A majority also agreed that negative public perceptions will slow or stop adoption of AI technologies. Many respondents were looking for internal governance mechanisms and more formal legal frameworks to manage the risks without stifling innovation.[13]

4.1 | AI governance as a pathway to responsible AI

💬 **The term "Responsible AI" has been embraced to describe a set of practices that ensures AI doesn't end up betraying people's trust in its application.**

With so much value at stake,[14] AI and the opportunity it represents remain at the top of the organization agendas, but they have also rightly reserved space for discussions over responsibility and trust. As concerns over bias, discrimination and privacy have proliferated in the news, industry leaders want to be sure to avoid negative consequences when they deploy AI.

The term "Responsible AI" has been embraced to describe a set of practices that ensures AI doesn't end up betraying people's trust in its application. Sitting at the core of many of those approaches are innovative governance mechanisms, which are quickly becoming a key competitive differentiator among companies seeking to adopt and leverage AI.

From big players to emerging start-ups, many companies recognize the potential for AI to do harm in the absence of safeguards and are proactively emphasizing the need to develop a strategy around AI that goes beyond legal risk management. Responsible AI aims to develop industry norms discouraging an AI ecosystem that is harmful to society or the planet.

Despite this growing awareness, translating intentions into policy and practice remains a challenge for companies. Among the biggest adoption issues for industry AI governance include a lack of executive understanding, practical and proven programs, and case studies. Industry wants to know more about what should be done, when and how to do it in practice.

## 4.2 | Risk-based approaches to identify and prioritize governance gaps
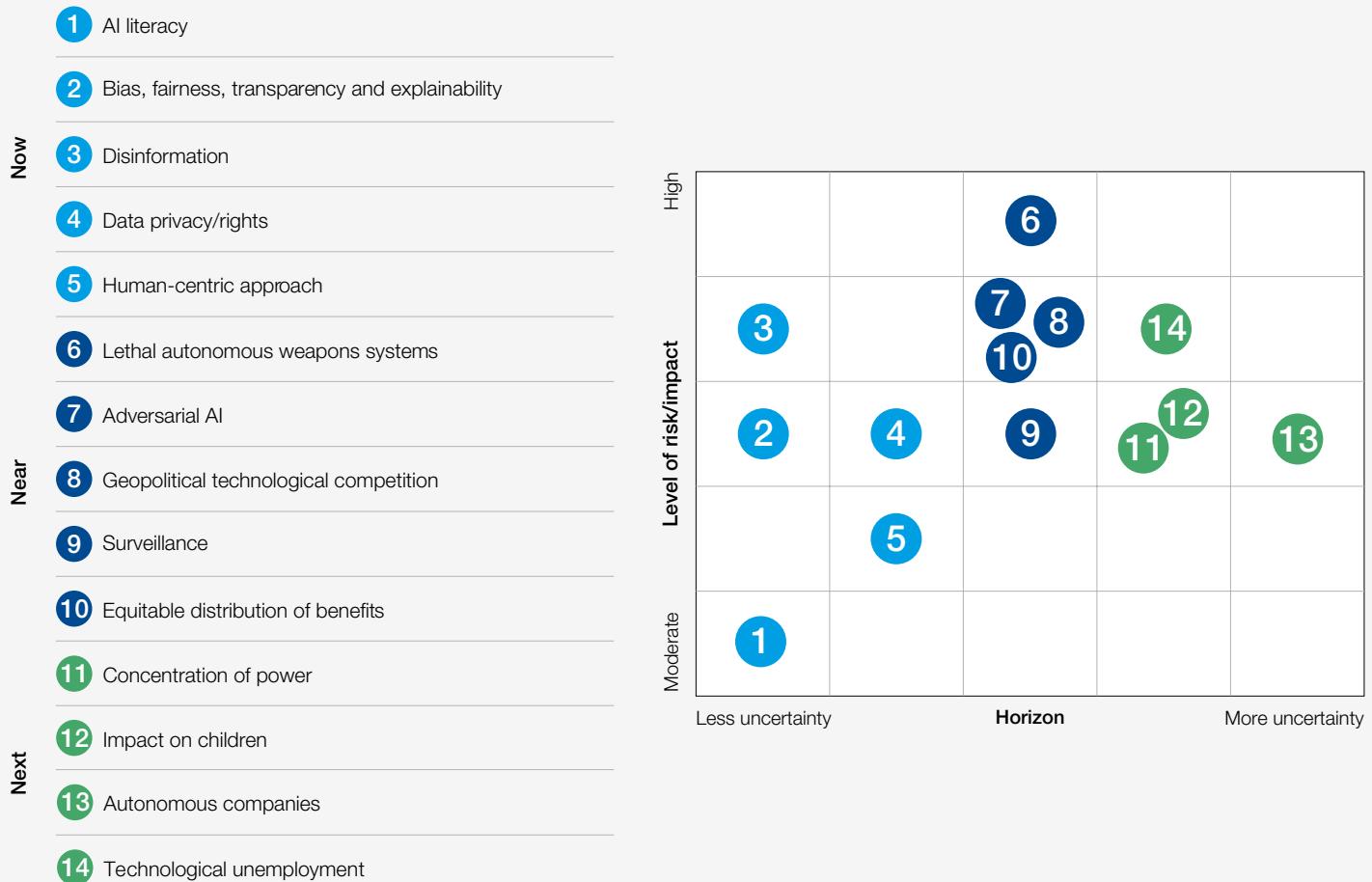
One of the greatest issues facing the private and public sectors alike is a lack of prioritization. Like other general-purpose technologies, AI promises to dramatically impact a range of different industries and sectors, but because its capabilities and implications are still being realized, the time horizon for its impact is less certain. The direct and indirect effects are expected to be felt at varying stages in the near to long-term future. This framing adds to the complexity of the discussion over how these systems should be governed. It can also help organizations strategize.

Taking the approach of examining the risks posed to society, organizations and individuals on a spectrum, many promising governance frameworks have begun to encourage the adoption of formal risk-management practices that help leaders identify and prioritize next steps. This can manifest in a number governance forms. For example, the German Data Ethics Commission proposed a five-level risk-based system of regulation, starting from no regulation for harmless AI systems to complete prohibition of the most dangerous ones.[15] Another example is Veritas, which is a part of Singapore's National AI Strategy and aims to provide financial institutions with a verifiable way of incorporating principles around fairness, ethics, accountability and transparency into their credit-risk scoring and customer marketing.[16]

Adopting frameworks facilitating better understanding of leaders about what risks are at play helps guide them towards first steps, whether they are as simple as knowing where AI-based systems are being applied or as committed as investing in a single executive in charge of AI risks. In a recent report, Deloitte offers an analysis of how many of the risks associated with AI and its existing or emerging governance gaps should be evaluated, placing them on a grid that measures them on two scales, risk level and time horizon (see Figure 2).[17] Such mapping can be helpful for leaders seeking clarity on the current governance outlook.

FIGURE 1 | Time horizon and risk level of emerging governance gaps

Now
1. AI literacy
2. Bias, fairness, transparency and explainability
3. Disinformation
4. Data privacy/rights
5. Human-centric approach
6. Lethal autonomous weapons systems

Near
7. Adversarial AI
8. Geopolitical technological competition
9. Surveillance
10. Equitable distribution of benefits
11. Concentration of power

Next
12. Impact on children
13. Autonomous companies
14. Technological unemployment

Source: https://www3.weforum.org/docs/WEF_Global_Technology_Governance_2020.pdf

4.3 | Promoting public knowledge and trust in AI

Over the past decade, the idea of making a push to increase consumer awareness around AI-based products has surfaced as a potential governance tool. This includes the use of certification, labelling and even awards to give AI purchasers a way of determining the extent to which an AI system is ethical and responsible. The rationale borrows from similar approaches taken in environmental policy and efforts to eradicate conflict diamonds.[18] The aim is to increase transparency around the product to influence purchasing behaviour, and through the power of markets, reputation and branding, the behaviour of AI-producing firms competing to earn client and user trust.[19]

Labelling and certification

Informative labelling schemes have been suggested as one way of signalling to consumers which companies handle data and AI in a trustworthy, ethical and secure way. Voluntary labelling schemes were also proposed in the European Commission (EC) white paper for AI applications that do not qualify as high-risk.[20] Examples include:

– **Data Ethics Seal (Denmark):** In 2019, Denmark along with a consortium consisting of the Confederation of Danish Industry, the Danish Chamber of Commerce, SMEdenmark and the Danish Consumer Council launched the prototype of a Data Ethics Seal. The seal represents an independent labelling scheme designating companies that meet its requirements for cybersecurity and responsible handling of AI-related data.

– **Labelling and specification frameworks (AEIG):** In 2020, the AI Ethics Impact Group (AEIG) released a framework aiming to support enforcement in Europe. It comprises three elements:

– First, an approach for the specification and operationalization of values;

– Second, a ratings system for AI ethics inspired by the energy efficiency label that would promote ethically sound AI systems in terms of transparency, grant users a standard for product comparison and provide policy-makers, regulators, standards developing associations with a basis for oversight along with surveillance organizations;

– Third, a risk matrix, a two-dimensional model for the classification of different application contexts of AI systems.

Certification is typically the process that leads to credible labelling schemes. It establishes criteria for qualification as well as an enforcement mechanism. Many of these are coming out of the private sector, but not exclusively. Its application has been proposed or applied to AI-systems in several instances. Examples include:

– **Ethics Certification Program for Autonomous and Intelligent Systems (IEEE):** Seeks to create specifications for certification and marking processes that advance transparency, accountability and the reduction of algorithmic bias in autonomous and intelligent systems.

– **Responsible AI Certification Program:** In collaboration with the World Economic Forum, AI Global and the University of Toronto (Canada) recently launched an initiative to develop a certification mark for the independent and authoritative assessment of AI systems.[21]

– **Voluntary AI certification programme (Malta):** In 2019, Malta became the first country in the world to launch an AI certification programme based on Malta's AI ethical framework.[22]

– **AI Certification (Singapore)**: Singapore's professional qualification programme for AI-related engineering roles includes a module on AI ethics, which tests candidates on the principles that underpin the importance of ensuring AI is built and used ethically, fairly and responsibly.[23]

Awards

By spotlighting and celebrating ethically and responsibly designed AI, award schemes can spread awareness and provide an incentive for industry to move in a positive direction. Example:

– **Smart Toy Awards:** The Smart Toy Awards, an initiative of the World Economic Forum, is a contest for companies that produce smart toys. Finalists, selected based on criteria co-developed with a multistakeholder community, were selected in April 2021 with winners announced during a virtual livestreamed awards ceremony in May 2021.

Algorithmic auditing

Third-party auditing can also be used to review algorithms and mitigate risk. External testing of complex systems has been effective in ensuring the health of financial markets, leading many to encourage the development of an industry focused on algorithmic auditing.[24] Applying the same level of objective scrutiny to AI ostensibly provides a check on AI, helping reduce the likelihood that faulty algorithms are released into the world.

The challenge before each of the schemes mentioned is in establishing credibility. For instance, algorithmic auditing requires the right incentives to ensure accountability.[25] One major criticism has been the lack of robust standards failing to hold the auditors themselves accountable in cases of a problem arising.[26]

To help overcome this, some have called for the professionalization of these efforts, which might entail the development of proper credentialing, standards of practice and disciplinary procedures.[27] With these more rigorous standards in place, methods such as auditing, certification and labelling could more effectively fulfill their promise of being an essential tool in operationalizing AI principles.

4.4 | AI governance by government

At the macro level, the pressure and urgency to develop effective rules and norms around AI is reflected in the number of efforts from the public sector – many of which have been tracked by the OECD Policy Observatory designed to share and help influence AI policies worldwide.[28] Some of these policies grew out of national AI strategies, including noteworthy legislative proposals, multinational partnerships and trade agreement provisions. Others included softer, norms-shaping approaches aiming to encourage the adoption of AI solutions in ways that both mitigate risks and accelerate the benefits.

Moves toward regulation

At the national and sub-national levels, there has been proposed legislation in the United States that would introduce mandatory impact assessments on high-risk automated decision-making systems[29] as well as outright bans on the use of facial-recognition technology (FRT) by law enforcement in numerous cities. At a multinational level, the EC has fully embraced the risk-based approach, acknowledging that the risks associated with AI systems are highly contextual. A 2021 proposal suggests stricter measures for high-risk AI applications, including safeguards against risks in sectors such as transportation and those that threaten individual rights in the EU human rights framework[30] (civil society groups have raised concerns that the approach did not focus sufficiently on protecting fundamental rights).[31]

The EC draft proposal classifies AI applications under four distinct risk categories:

1. **Unacceptable risk**: AI applications in this category will be banned (e.g. social scoring)

2. **High risk**: AI applications will be subject to a targeted vetting process which includes quality management and conformity assessment procedures (e.g. AI recruitment tools/ facial recognition)

3. **Limited risk**: such applications will be subject to minimal transparency obligations (e.g. chatbots)

4. **Minimal risk**: applications in this category will not be subject to any additional provision (e.g. spam filters)

The EC proposal is designed to target specific use cases. Instances like university admission, hiring for employment, access to financial credit and the use of facial-recognition technology to identify individuals are to be stringently regulated. This will involve deploying risk-mitigation processes such as third-party certification for technology providers. Where no specific threat from AI systems has been identified, minimal regulation is recommended to encourage innovation in the European Union.

Beyond the EU, AI governance provisions are also inscribed in multilateral trade agreements such as the US-Mexico-Canada Agreement (USMCA, effective 1 July 2020, replacing NAFTA) as well as the Digital Economy Partnership Agreement between Singapore, New Zealand and Chile and the Australia-Singapore Digital Economy Agreement.[32]

Important milestones in government action towards AI governance:

– March 2017: Canada becomes the first country to adopt a national AI strategy

– July 2017: Japan publishes Draft AI R&D Guidelines for International Discussions, proposing a set of non-binding AI R&D principles

– June 2019: Ministry of Economy, Trade and Industry (METI) of Japan publishes the Contract Guidelines on Utilization of AI and Data

– September 2019: United Kingdom Office for AI releases a draft set of guidelines for AI procurement to support government agencies in their responsible purchase of AI

– January 2020: Singapore introduces the second edition of its Model AI Governance Framework to help the private sector with recommended AI governance mechanisms

– February 2020: European Commission releases a White Paper on AI detailing its comprehensive approach on these issues

– June 2020: Global Partnership for AI (GPAI) is launched

– April 2021: European Commission releases proposed regulations for AI

Accelerating international cooperation

Simultaneously, governments have led on several promising, soft AI governance mechanisms. Recognizing that leaving jurisdictions to govern unilaterally would lead to a global patchwork of fragmented approaches that ultimately undermine the potential of AI, many governments have pursued efforts that prioritize fostering international cooperation and harmony of ethical guidelines across national boundaries.

Regional powerhouses have been particularly active in their attempts to establish some semblance of a united front, including efforts stemming from the G-7, which set out in 2017 to create a unified body encouraging a common understanding and approach to AI. This led to the establishment of the Global Partnership for AI (GPAI) in 2020, a multistakeholder initiative comprising 18 countries and the EU.[33] The 37-member OECD serves as secretariat for GPAI. The G-20, which includes China, the Russian Federation and Saudi Arabia, has not moved beyond the adoption of a set of AI principles at the 2019 Osaka G-20 meeting.[34]

At the global level, the United Nations have taken a leading role in balancing technological progress and social progress, with the International Telecommunication Union (ITU) and its AI for Good Global Summit at the forefront. Other UN agencies have also contributed to AI governance efforts by promoting research and encouraging dialogue, including International Labour Organization work on AI impact on work and jobs and the UNESCO Global Dialogue on the Ethics of AI.[35] UNICEF's work, *Policy Guidance on AI for Children*, which partners with the World Economic Forum Generation AI project, highlights the need for not just human-centric AI but child-centric AI.[36] In doing so, it aims to encourage governance mechanisms that accelerate the benefits that children can gain from AI advancements while also pointing to the unique risks faced by children.

Greater international cooperation and an emerging consensus on principles has been a positive sign in recent years. To move from principle to policy to practice, however, the challenge now is to arrive at more consensus and clarity in terms of governance approaches and operational guidance. One example of how this might be done comes from the UK, which in 2020 published guidance on what it means to make AI explainable. Under Article 22 of the General Data Protection Regulation (GDPR), systems that are fully automated with legal impact on consumers (e.g. loan decisions, recruitment, etc.) need to be explainable.[37] To help guide companies, the UK Information Commissioner's Office offers detailed operational guidance to companies.[38] Another example comes from the Personal Data Protection Commission (PDPC) of Singapore, which published a *Model Artificial Intelligence Governance Framework* that provides detailed and implementable guidance for private-sector organizations to address key ethical and governance issues when deploying AI solutions.[39]

As governments move forward with new regulations, it will also be important to avoid conflicting obligations in different jurisdictions that inhibit cross-border research and impede innovation. Though specific regulations will certainly differ, the rules that are eventually codified will need to remain as interoperable and aligned as possible.

BOX 1	COVID-19

Even before the COVID-19 pandemic broke out, AI governance was an urgent issue. In a matter of months, the crisis dramatically accelerated business demands for AI solutions while also driving uncertainty even further.
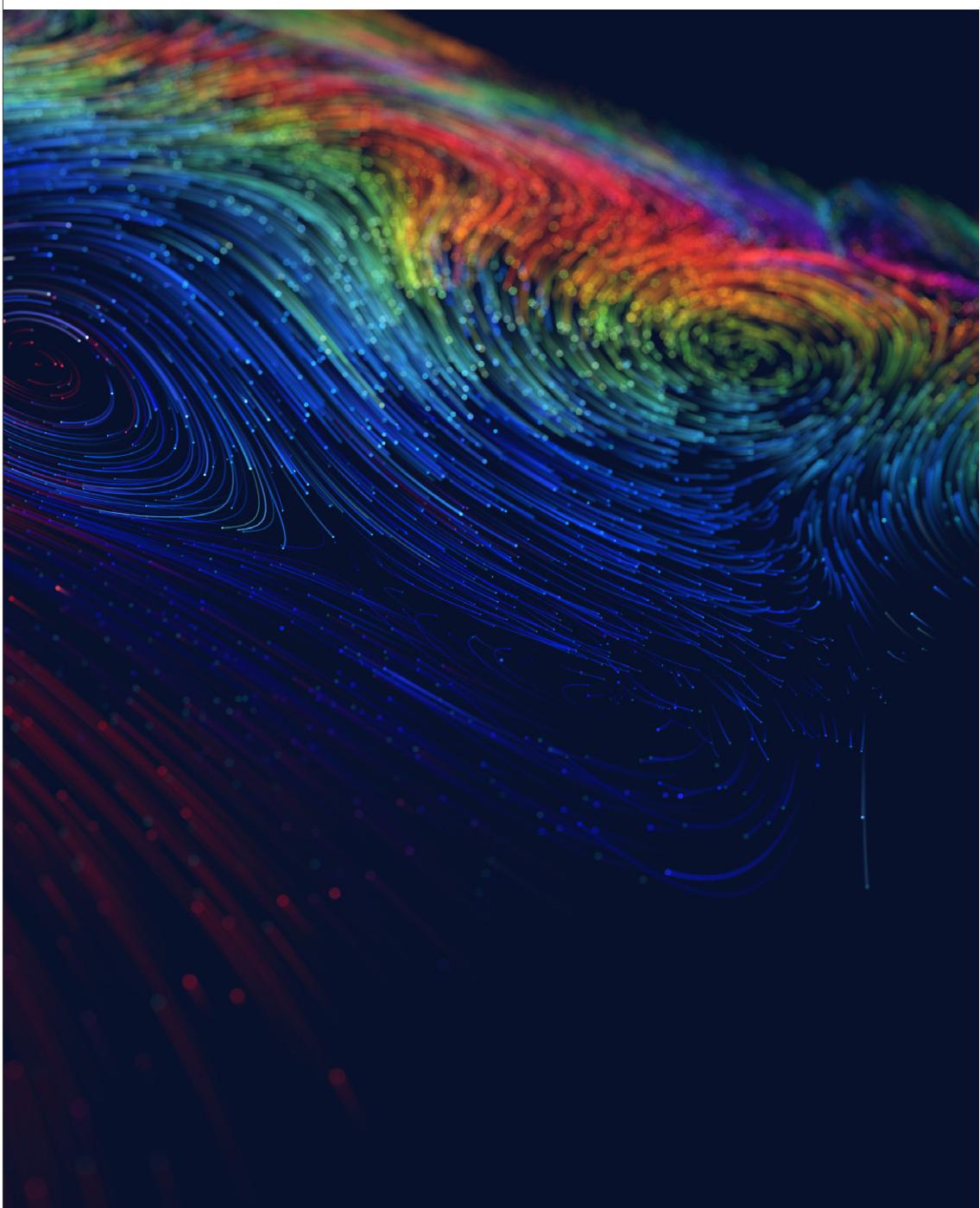
On the one hand, the technology was able to play a major role in creating the efficiency and continuity that companies and governments desired as they pivoted to remote or contactless situations. Conversational AI, also known as Chatbots, became a natural partner in disseminating health information during the coronavirus crisis. Countries began using AI-driven contact tracing algorithms to send citizens personalized text messages, instructing them to isolate after being near someone with a positive diagnosis.[40] The need for seamless and contactless technology to accurately identify customers, employees and vendors led to greater adoption of remote biometric AI systems such as FRT.[41]

On the other hand, the pandemic served to highlight what held AI back. Despite perhaps a higher willingness among populations to trust emerging technology in the midst of a crisis, initial announcements to deploy contact-tracing apps sometimes triggered public outcries over privacy concerns.[42] Chatbots that retrieved information from back-end repositories that were not updated frequently provided inconsistent results.[43] Lacking large, high-quality and accurate data sets, most AI-powered contact tracing interventions fell flat.[44]

Much of the impediment resulted from a lack of data as well as fears over an erosion of data privacy. Once the pandemic subsided, uncertainty around whether measures to survey populations would be rolled back has been cited as a significant limiting factor.[45] These constraints made the case for effective AI governance even stronger.

5 Multistakeholder approaches

Industry, government, academia and civil society have all been instrumental in the development of AI governance

Led by academia, advocacy organizations, think tanks and international organizations such as the United Nations, the OECD and the World Economic Forum, many of the most promising AI governance frameworks have emerged from initiatives that convene a variety of actors. The Forum's Global AI Council has championed this approach.

Given the cross-cutting nature of AI systems, stakeholders working in silos inevitably run up against a range of challenges, including barriers to cooperation, knowledge gaps, divergent interests and conflicting incentives. Enhanced multistakeholder approaches on the other hand have helped build global capacity for the development of responsible AI by filling knowledge gaps and ensuring efforts are more human-centred. Going at it alone does not work in AI governance.

Using a multistakeholder approach also helps organizations gain social license from the public and in the process of in doing so, they earn trust. Both the progress achieved and the obstacles confronted over the past five years reinforce the need to build trust in AI systems.

Today, several examples demonstrate how trust can be earned by building constructive relationships with stakeholders across sectors, regions and other categories to together craft and sustain responsible AI governance mechanisms. These relationships require open and honest conversations with the people who will be affected by AI, whether they are customers or citizens. In turn, social licence is granted to organizations that are deemed trustworthy even when serious risks are involved.

5.1 | Stakeholder involvement

The key question in these efforts is how and what various actors can contribute. Each stakeholder brings a vital perspective and each group has a vested interest in ensuring responsible governance of AI.

Industry

There is a tremendous opportunity for industry to build upon and scale innovative approaches to AI governance at this critical juncture by sharing global best practices for trustworthy AI. The reasons for engaging go beyond anticipating future regulation or merely managing legal and reputational risk. There are commercial benefits as well. According to a recent jointly published survey on AI use and adoption by the Boston Consulting Group and MIT, 90% of organizations are developing or have already developed strategies aiming to improving fairness in algorithms and reducing bias in decision-making. Of those, 72% find that these strategies increase the financial benefits of AI, and 62% report that they decrease operational risk.[46]

Steps that companies can take to support multistakeholder efforts include:

– **Participate in a "regulatory sandbox"**: The growing need to develop governance frameworks for the use of AI has highlighted the need to test and experiment in controlled environments. These piloting opportunities provide untested frameworks with a dose of the real world and give companies a chance shape the future of AI governance.

– **Offer technical or executive expertise**: Engage with multistakeholder efforts by seconding personnel who can fill knowledge gaps and inform the design of governance frameworks.

– **Donate technical tools that bolster their AI governance**: As the primary holders of AI expertise, many companies are beginning to develop and offer tools that help others evaluate their AI models. Examples include:

  – The AI Fairness 360 Toolkit: An open-source library for data scientists that allows AI model builders to identify, investigate and mitigate unwanted bias in their models.

  – Watson OpenScale: An open platform designed by IBM Watson that helps businesses explain AI models and outcomes to business users, serving as a "one-stop-shop" for monitoring fairness within the context of other issues such as explainability, accuracy and model health.

  – Google What-if tool (WIT): Helps enable people to examine, evaluate and compare machine-learning models.

Government

As governments strive to leverage AI for better services, they first require social license. In addition, as they consider moving forward with regulations, it will also be important to avoid conflicting obligations in different jurisdictions that inhibit cross-border research and impede innovation.

Steps that governments can take to support multistakeholder efforts include:

– **Develop "regulatory sandboxes"**: Policies need places where new governance ideas can be quickly deployed, tested and iterated. Governments can play a critical role in ensuring that AI is tested in as real an environment as possible before being released to the world.

– **Offer public policy or executive expertise**: Engage with multistakeholder efforts by seconding personnel who can fill knowledge gaps, especially around public policy and can offer perspectives on the unique pressures faced by government actors.

– **Set up a centre of excellence for AI**: Establishing a national or subnational body focused on AI governance can help draw staff from industry, government, academia and civil society, using a multidisciplinary and collaborative approach to provide advice on AI and algorithm use for government operations. Countries that have so engaged include India, Malta, New Zealand, Singapore and the UK. [47]

Academia

In addition to bringing crucial scholarship to the table, academia is also uniquely positioned to operate independently at the intersection of sectors, helping provide a level of neutrality and objective perspectives that can be effective in attracting others to the conversation.

Steps academia can take to support multistakeholder efforts include:

– **Provide opportunities for research:** As policies are developed and tested, academia can conduct impact assessments measuring the effectiveness of new approaches.

– **Foster an interdisciplinary approach to responsible AI:** Studying the ethical implications of AI requires a variety of disciplines across the humanities as well as social and formal sciences to come together to contribute to the AI discussion.

– **Help disseminate and scale promising AI governance mechanisms:** Academia can be vital partners in helping ensure practices and behaviors informed by the multistakeholder process make it into curricula around the world.

Civil Society

The role of civil society in the development of responsible AI is important because it represents a third and crucial voice in discussions over how the technology can benefit all parts of society. Civil society can bring balance and perspective to the conversation. It can fill blind spots and put fears at ease among those who stand to lose the most in a world without proper guardrails for AI.

Some of the steps civil society can take to support multistakeholder efforts include:

– **Close participation gaps**: Civil society has long helped advocate on behalf of vulnerable or marginalized populations. In the context of AI, this could mean encouraging more diverse and representative data sets. It could also

entail supporting AI leaders from the global south to participate in international discussions on AI governance.

– **Convening diverse actors and building community**: Like academia, civil society, too, can take on the role of convener for a range of different actors.

– **Engage with centres of excellence for AI**: In jurisdictions that create bodies mandated to advise on the use of AI, especially when it is being applied on the public, civil society can organize and encourage broad engagement with these entities to ensure policies are informed by a diverse and inclusive discourse.

BOX 2 | Agile governance

Agile governance has been heralded as a way in which governance can keep pace with exponential technological advancement and its impact on society. The methodology requires timely experimentation and decision-making without sacrificing the duty of public and private governance processes to empower and protect those they serve. A full description of the agile methodology can be found in the World Economic Forum white paper on *Agile Governance: Reimagining Policy-making in the Fourth Industrial Revolution*.[48] In collaboration with the Governments of Canada, Denmark, Italy, Japan, Singapore, the UAE and the UK, the World Economic Forum launched the Agile Regulation for the Fourth Industrial Revolution project which recently produced a guide on strategies to respond in a more agile way to innovation and disruption, building on the latest evidence and practice from around the world.[49]
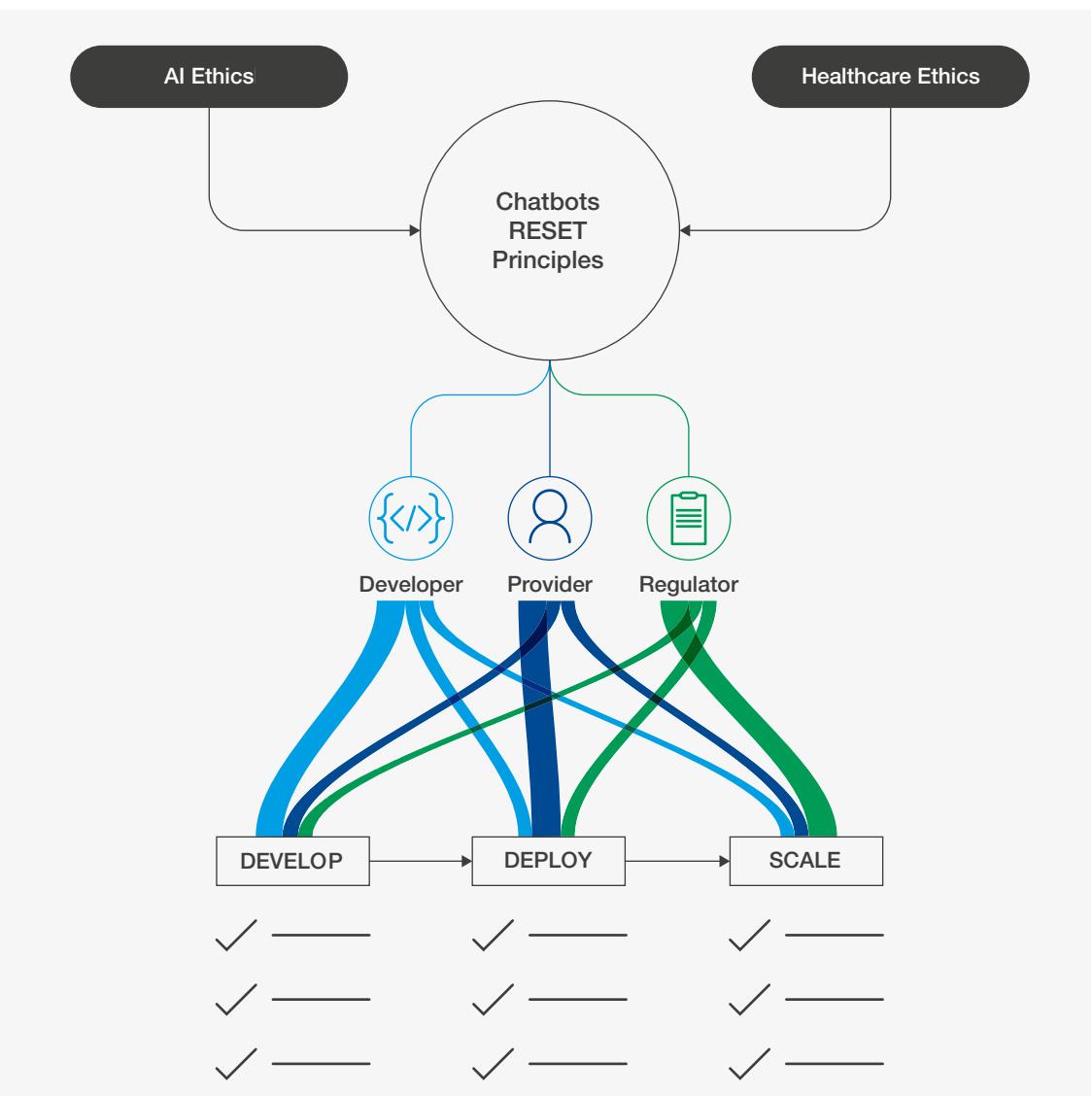
5.2 Agile governance

When you combine the multistakeholder approach with an agile governance methodology, you enable policy-making that is more inclusive and human-centred by involving more people in the process and allowing for rapid iteration to meet the needs of everyone involved. Below are some examples of multistakeholder communities that drew on existing, generalized principles to design and test a framework tailored to a specific use case in a given context. Each was advised and supported by the Forum's Global AI Council.

FIGURE 2 | **Chatbots report**

Source: *Chatbots RESET A Framework for Governing Responsible Use of Conversational AI in Healthcare*, https://www3.weforum.org/docs/WEF_Governance_of_Chatbots_in_Healthcare_2020.pdf

- **Chatbots in healthcare**: After seeing a tremendous uptake in the use of chatbots during COVID-19, Mitsubishi Chemical Holdings, Reliance Industries, dozens of other companies and the Government of India worked with the World Economic Forum to quickly design frameworks for governing chatbots in healthcare by bringing together chatbot developers, chatbot platforms, the medical community, civil society, academia and healthcare regulators. There are presently five pilots underway with a few more being planned.[50]

- **Facial recognition in flow management and law enforcement**: This community aims to create a governance framework to address concerns arising specifically from the use of FRT for flow management at heavily-populated locations such as airports, and for law enforcement. In 2020, the project led to the world's first publicly shared FRT self-assessment by an organization after being piloted by Narita International Airport Corporation in Japan. It also released a certification scheme to ensure the responsible use of FRT for flow management.[51]

- **AI Procurement guidelines for government**: Hoping to leverage the government's purchasing power rather than solely relying on its regulatory power, the UK Office for AI developed a set of guidelines in collaboration with the World Economic Forum, the Government Digital Service (GDS), Government Commercial Function and Crown Commercial Service.[52] The guidelines have since been piloted with seven public-sector entities across five jurisdictions. Three to five additional pilots are planned for 2021.

- **Model AI governance framework**: In 2019, the Government of Singapore's Personal Data Protection Commission (PDPC) released the second edition of its *Model AI Governance Framework*, a sector-, technology- and algorithm-agnostic framework that converts relevant ethical principles to practices in an AI deployment process so that organizations can operationalize these principles. Singapore then designed a companion *Assessment Guide* that was released in 2020 to help with implementation.[53]

- **Reimagining Regulation for the Age of AI**: In the first global multistakeholder effort to co-design regulatory frameworks for AI informed by a policy pilot, the Government of New Zealand has for the past year and half been working to build social license for the use technology. This required the involvement of stakeholders from industry, academia and nationwide discussions with New Zealand's indigenous people, the Māori.[54]

- **Human-Centred AI for Human Resources**: This project is developing a toolkit for human resource professionals to promote the human-centred and ethical use of AI. The community is currently working with select industry, government and professional association partners to pilot toolkit and revise.[55]

5.3 | Global AI Action Alliance

Although much has been done to help address the issue by bolstering international coordination among governments, less has been done to date to help address the growing proliferation of private or internal governance mechanisms at a global level. That is, in part, what inspired the Global AI Council in May 2020 to endorse the development of an accelerator to help drive global collaboration and rapid scaling of proven tools and practices for AI governance.
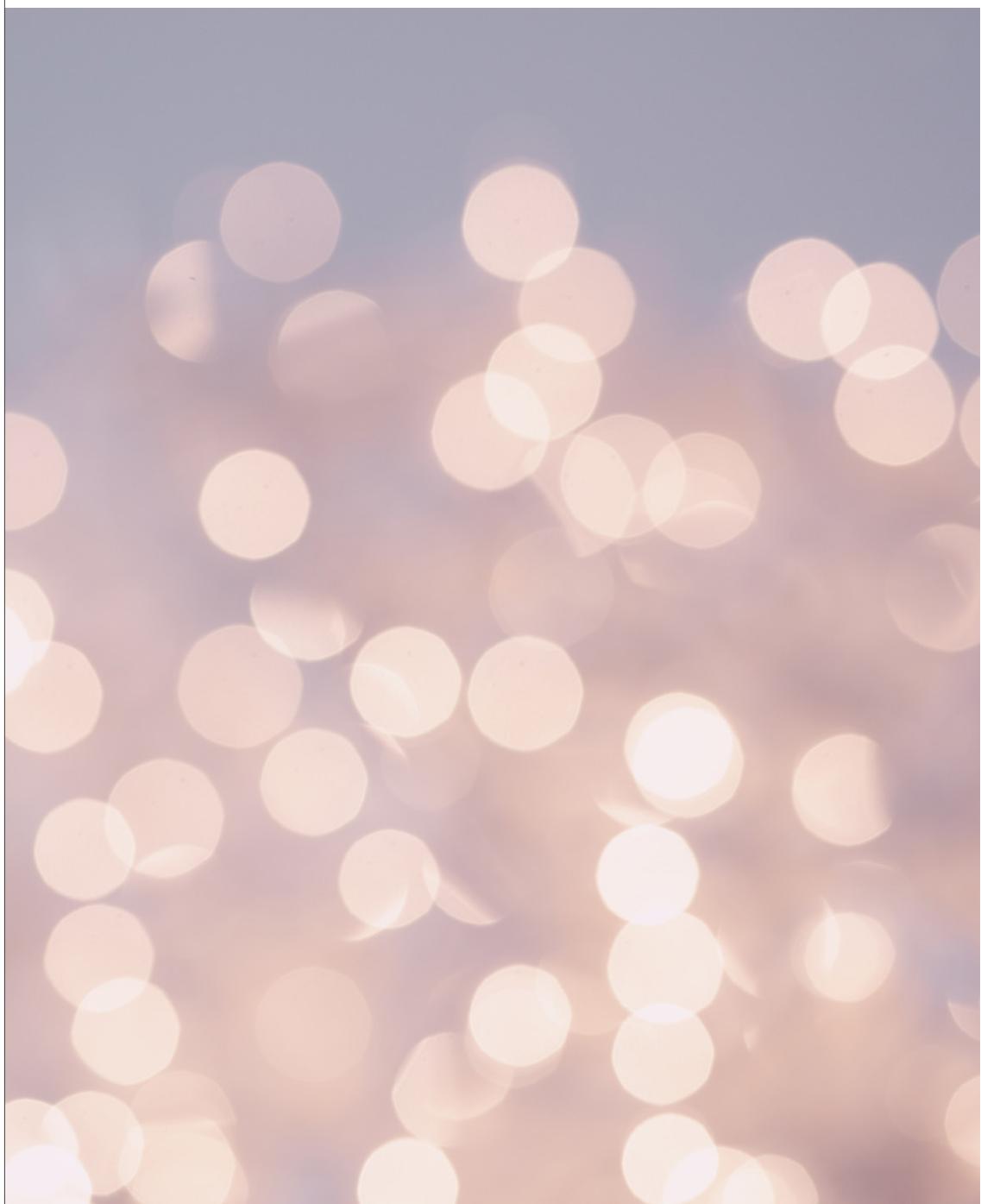
In turn, the World Economic Forum launched the Global AI Action Alliance (GAIA) in January 2021. The Patrick J. McGovern Foundation, a funder committed to data and AI for good, is supporting the alliance as part of a larger $40 million commitment. GAIA builds on the Forum's global multistakeholder community of leading businesses, governments and civil society organizations engaging actively with AI, and over three years of successful efforts to create interoperable governance mechanisms for the development and use of AI technologies. GAIA will also collaborate closely with the Forum's Industry Action Groups and its Data for Common Purpose Initiative (DCPI) to ensure that AI governance reflects best practices in data governance and is interoperable globally and across industry sectors.

⑥ The Road Ahead

Experts highlight governance gaps and use cases that will be key in the coming years

6.1 | More tools and best practices

Although we are getting smarter about where the gaps exist, for whom and on what timeline, this analysis changes depending on industry, region and social class. And while AI governance has been buttressed by expanded points of intervention, there remains a need for additional tools and practices. Even as organizations are increasingly aware the AI risks operate on a spectrum, there is now a need to clearly define where the lines should be drawn and in what contexts. Where can we rely on self-regulation, standard contract provisions and where is affirmative regulation necessary?

The Global AI Council recommends that the following issues warrant more attention in the year ahead:

– Standards providing a framework for responsible AI (e.g. What are the dimensions of a responsible AI system? What criteria defines a high-risk use case?)

– Standards for measuring bias, fairness and related technical details

– Processes and tools for assessing AI systems (e.g. questionnaires)

– Tools for increasing executive awareness, understanding and prioritization

– Approaches to cross-cutting ethical issues that will require public intervention and alignment

(e.g. disinformation, deep fakes, large-scale facial recognition, AI-enabled surveillance)

– Case studies in industry operationalization of AI governance that provide industry examples of how to do it practice

– Incentives that not only help address research gaps in AI but also help encourage AI researchers to think carefully about the broader impact of their work and to choose research directions accordingly. A recent paper offers a balanced discussion of the pros and cons of this approach, along with suggestions for further improvements.[56]

The Council also encourages leveraging learnings from the experience of governance of other high-risk technologies. For instance, the Asilomar Conference on Recombinant DNA serves as a good example of the private sector measuring risk and then cooperating with the public sector through grant review to certify and protect against negative outcomes.

Other examples include:

– International Panel on Climate Change[57]

– Flu research panel[58]

– Asilomar for CRISPR-CAS9[59]

6.2 | AI, jobs and inequality

Largely unaddressed by this report but important to the Global AI Council are issues around the impact of AI on jobs and inequality. AI is impacting and will continue to impact economies around the world. One of the challenges concerns the millions of people who are contracted to collect and label the data used in machine-learning models. This global workforce, sometimes referred to as the "invisible workers of AI," is left largely in the shadows. Ensuring their ethical treatment requires more attention.

Another challenge concerns accelerated task encroachments. At the time of writing, most AI systems have narrow specialization: they are very good at one single function. As a result, many physical tasks previously considered un-automatable can now be performed by machines, from medical diagnoses to legal document drafting. Such trends are fueling the growth of automation technologies and will spur large-scale, global changes, including 400 to 800 million lost jobs by 2030, according to a 2017 McKinsey report.[60]

This is concerning enough, but as these capabilities expand, so too will the issues around inequality. AI could substitute certain types of tasks and occupations leaving some people unemployable while others can still sell their skills to perform tasks not yet automated by AI, which will serve to concentrate wealth if better policies around distribution and reskilling are not implemented. Even very optimistic economists anticipate a world that promises a high standard of living to all but leaves few people with any useful role that is valued by others by today's standards, which might trigger a crisis of meaning in many people's lives. It is also not unthinkable that one or more major company may one day find itself in control of a substantial fraction of the world's economy.

In order to avoid such outcomes, suggested policy levers have included the so-called Windfall Clause[61], where AI firms commit to donating, in such a situation, a significant share of their profits to the common good and the broad and equitable

benefit of humanity. To help explore policy responses further, the Forum's Global AI Council is pursuing an ambitious initiative called Positive AI Economic Futures, taking as its starting point the hypothesis that AI systems will eventually be able to do the great majority of what we currently call work, including all forms of routine physical and mental labour. In a series of ongoing workshops, they have brought together a diverse group of individuals to discuss existing visions and their implications for present-day policy.

6.3 | AI as an emitter of carbon

As knowledgeable as we have become in tackling some areas, a considerable amount of thought and work remains on other downstream effects of AI on the planet. Though certainly championed for its potential to help tackle global issues such as climate change,[62] the infrastructure around AI systems has also come under scrutiny for its carbon output.[63]

6.4 | Future gaps

Like AI, quantum computing promises to help provide significant advances in computing capabilities and will have considerable impact on improving the state of the world. But as with AI, along with the emergence of a global quantum ecosystem the inherent risks of reliance on quantum computing have also emerged across a number of areas.

Will quantum computing add to the risks, as well as to the benefits, of an algorithm-based society? Are there new challenges? Will the public trust technologies that they cannot understand and whose results they cannot verify? Will quantum computers be able to break existing cryptographic standards? Who will have access to the new technologies?

Discussions around these questions have begun under the auspices of the Global AI Council.

BOX 3 | **Quantum – Why should we be excited?**
Nikhil Malhotra, Chief Innovation Officer, Tech Mahindra
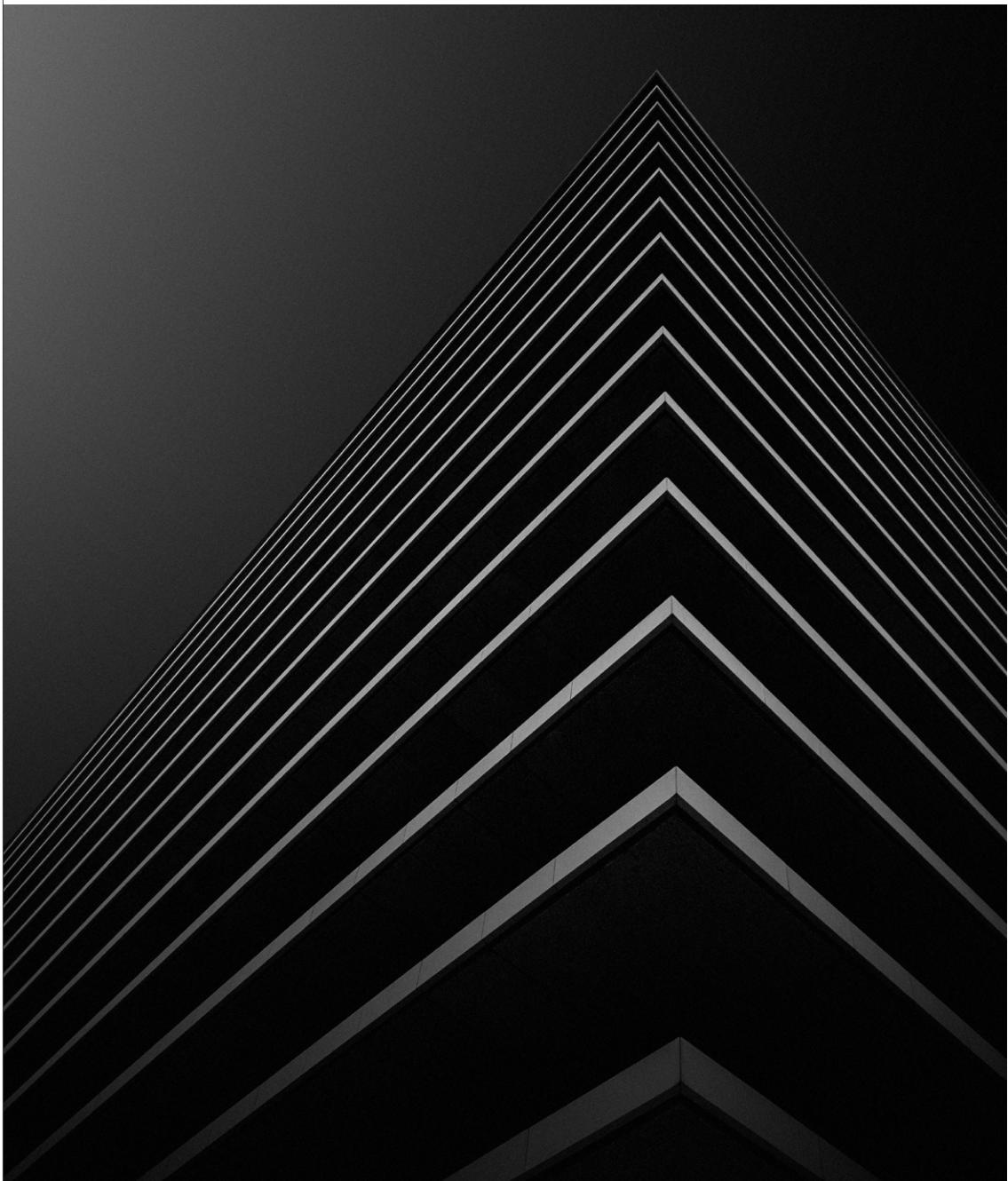
We should be excited about this subatomic world for the following reasons:

– It is no longer science fiction and practical cases of quantum computing have started emerging

– Edges of classical computing would be reached very quickly. If we look at the size of the electronics today, at around 0.5 NM scale of VLSI design, electron tunneling would start playing havoc with the consistency and parity of information

– The current global dependency on energy for technology has become unsustainable so a fundamental shift in the mechanism of computing is needed. The primary reason for that is that classical computation does an irreversible computing paradigm and quantum computing follows a reversible principle. Reversible cycles or computation implies that, given the output and the operator, we can get back the input which makes the computation reversible

– The complex problems or NP (Non deterministic polynomial) hard problems as they are referred to feature a facet that can be resolved using quantum AI algorithms

  – Optimization

  – Protein folding and chemical analysis can now be handled using a quantum computer. Until the onset of quantum, the chemical formula that a machine could understand well was Caffeine (Chai). The reason is the covalent bond in a molecule and each covalent bond lends itself to a million electrostatic or ionic transactions between electrons that is hard even for supercomputers to handle

Quantum computing would help us boost the power of machines and AI by thinking at a fundamental level of universe and applying. Be it augmented experiences in retail, travel and hospitality, which has become quintessential during COVID-19 times, or ensuring that an epidemic like this in the future is predicted and acted upon at a rapid pace, AI, quantum and other deep technologies would become an essential tool in a country's arsenal.

(7) Conclusion

The journey to construct an ecosystem of oversight has accelerated considerably in the past five years as organizations around the world, including entities from the public and private sectors, have scrambled to close the governance gaps opening up as AI acceleration has outpaced traditional rules and regulations. As more automation enters all aspects of our daily lives, it is important to have ways to oversee the operations of these systems to ensure they are helping, not harming, humanity. The flurry of governance frameworks over the past two years has been crucial in helping leaders to better understand the issues surrounding AI, including potential for fairness and discrimination, disparate impact and the associated issues of transparency and accountability. But much more innovation in the realm of AI governance is needed if we are to keep pace with both the advancement and application of AI-based systems. To accomplish this, we need to double down on multistakeholder efforts. By bringing multiple actors together to co-design and co-develop governance frameworks, this model acknowledges that the whole is greater than the sum of its parts.

Contributors

This insight report is a product of multistakeholder collaboration and consultation. The content herein is the culmination of research, workshops, interviews, feedback and meetings that the Global AI Council has generously contributed over the past two years. These leaders play a critical role in providing a diverse set of perspectives on the various approaches to AI governance.

Members of the Global AI Council

Omar bin Sultan Al Olama
Minister of State for Artificial Intelligence, Digital Economy and Remote Work Applications, Office of the Prime Minister, United Arab Emirates

Sam Altman
Chief Executive Officer, OpenAI, USA

Irakli Beridze
Head, Centre for Artificial Intelligence and Robotics, UNICRI, Netherlands

Mark Brayan
Chief Executive Officer, Appen, Australia

Paul Daugherty
Chief Technology Officer, Accenture, USA

Umeshwar Dayal
Senior Vice-President and Senior Fellow, Hitachi America, USA

Virginia Dignum
Professor, Ethical and Social Artificial Intelligence, Umeå University, Sweden

Caroline Dinenage
Minister of State for Digital and Culture, Department for Digital, Culture, Media and Sport, United Kingdom

Carmine Di Sibio
Global Chairman and Chief Executive Officer, EY, USA

Erin Finkel
Associate Director, Brand, Marketing and Communications, EY, USA

Marine Formentini
Chief of Staff, Alpha Intelligence Capital, France

Nicole Foster
Head, Global AI & Head, Public Policy, Canada, Amazon Web Services, Canada

Simon Greenman
Founding Partner, BestPractice.AI, United Kingdom

C.P. Gurnani
Managing Director and Chief Executive Officer, Tech Mahindra, India

Michael Harasimowicz
AI Engineer, Lockheed Martin, USA

Matissa Hollister
Assistant Professor of Organizational Behaviour, McGill University, Canada

Jeremy Howard
Founding Researcher, fast.ai, USA

Jas Jaaj
Managing Partner, Artificial Intelligence and Data, Deloitte, Canada

Santeri Kangas
Chief Technical Officer, CUJO AI, Finland

Amitabh Kant
Chief Executive Officer, NITI Aayog, India

Konstantinos Karachalios
Managing Director, Institute of Electrical and Electronics Engineers (IEEE) Standards Association, USA

Fayaz King
Deputy Executive Director, Field Results and Innovation, UNICEF, USA

Andreas Kunze
Chief Executive Officer and Co-Founder, KONUX, Germany

Fei-Fei Li
Sequoia Professor, Co-Director of Stanford's Human-Centered AI Institute, Stanford University, USA

Jesus Mantas
Senior Managing Partner, Global Strategy, Innovation and Corporate Development, IBM Global Business Services, IBM, USA

Steven Mills
Chief Artificial Intelligence Ethics Officer, Boston Consulting Group (BCG), USA

Illah Nourbakhsh
Professor, Robotics Institute, Carnegie Mellon University, USA

Michael Priddis
Chief Executive Officer, Faethm.ai, Australia

Endnotes

1. https://www.accenture.com/_acnmedia/pdf-24/accenture-universal-principles-data-ethics.pdf.

2. Smith, Brad, and Browne, Carol Anne. *Tools and Weapons*. p. 195.

3. Schwab, Klaus. *Shaping the Future of the Fourth Industrial Revolution.* P. 123.

4. https://techcrunch.com/2014/01/26/google-deepmind/.

5. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

6. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

7. https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF.

8. Hagendorff, T. (2020). "The Ethics of AI Ethics: An Evaluation of Guidelines." https://doi.org/10.1007/s11023-020-09517-8.

9. Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy and Madhulika Srikumar. *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI* (15 January 2020). Berkman Klein Center Research Publication No. 2020-1, available at SSRN: https://ssrn.com/abstract=3518482 or http://dx.doi.org/10.2139/ssrn.3518482.

10. https://urldefense.proofpoint.com/v2/url?u=https-3A__openai.com_blog_ai-2Dand-2Dcompute_&d=DwMFaQ&c=VWART3hH1Kkv_uOe9JqhCg&r=sKWjdT1dnJbq7vgTHoQUVgM3vMFjx8xgv1S2Xtep0Sc&m=sOsSqCVZNx-pmvfg-TRlxyHIM7CYonOBYsPZ7tUAdU0&s=EhbDa2nfsK_0fCoOO6fKe1SNqOXTJtONpaHBG1kHOso&e=.

11. https://urldefense.proofpoint.com/v2/url?u=https-3A__openai.com_blog_ai-2Dand-2Defficiency_&d=DwMFaQ&c=VWART3hH1Kkv_uOe9JqhCg&r=sKWjdT1dnJbq7vgTHoQUVgM3vMFjx8xgv1S2Xtep0Sc&m=sOsSqCVZNx-pmvfg-TRlxyHIM7CYonOBYsPZ7tUAdU0&s=3izqDAA7KElhOIBtE8m8zvRkg6OZwQ2hTf1YHKvO2dI&e=.

12. https://cyber.harvard.edu/publication/2020/principled-ai.

13. https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/state-of-ai-and-intelligent-automation-in-business-survey.html.

14. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-executives-ai-playbook?page=industries/.

15. https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html.

16. https://www.mas.gov.sg/news/media-releases/2021/veritas-initiative-addresses-implementation-challenges.

17. World Economic Forum (2020). *Global Technology Governance Report 2021: Harnessing Fourth Industrial Revolution Technologies in a COVID-19 World*, http://www3.weforum.org/docs/WEF_Global_Technology_Governance_2020.pdf.

18. https://www.kimberleyprocess.com/.

19. https://doi.org/10.2307/1879431.

20. "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS", 2016, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206.

21. https://ai-global.org/certification/.

22. https://www.maltachamber.org.mt/en/malta-first-country-in-the-world-to-launch-ai-certification-programme

23. https://www.aisingapore.org/ai-certification/.

24. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf.

25. https://www.brookings.edu/research/auditing-employment-algorithms-for-discrimination/.

26. https://www.technologyreview.com/2021/02/11/1017955/auditors-testing-ai-hiring-algorithms-bias-big-questions-remain/.

27. https://hbr.org/2018/11/why-we-need-to-audit-algorithms.

28. https://oecd.ai/.

29. https://www.govtrack.us/congress/bills/116/s1108.

30. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

31. https://www.accessnow.org/eu-regulation-ai-risk-based-approach/.

32. https://www.brookings.edu/research/strengthening-international-cooperation-on-artificial-intelligence/.

33.	https://gpai.ai/.
34.	https://www.oecd.org/about/secretary-general/2019-g20-leaders-summit-digital-osaka-june-2019.htm.
35.	http://handle.itu.int/11.1002/pub/813bb49e-en.
36.	https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children.
37.	https://www.privacy-regulation.eu/en/article-22-automated-individual-decision-making-including-profiling-GDPR.htm.
38.	https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence-1-0.pdf.
39.	https://www.pdpc.gov.sg/Resources/Model-AI-Gov.
40.	https://techcrunch.com/2020/03/18/israel-passes-emergency-law-to-use-mobile-data-for-covid-19-contact-tracing/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAANbmErCyLYW-o1XzERNinlRDaHcjJz-Sgx3WjR8-oMgimV9SXa1gQlvUTy9CPZ6a-PxjoQ1Ug8W3HyjCVUD1fhJxNvsvENkcp4dRefv0xm9rUFBapzaPv-2iFwj0Vh2Tziiz7bpf_dHVzrZlgKl0FURY1TCEKJResH0hQqb5aBXT.
41.	http://www3.weforum.org/docs/WEF_Responsible_Limits_on_Facial_Recognition_2020.pdf.
42.	https://www.zdnet.com/article/singapores-move-to-introduce-wearable-devices-for-contact-tracing-sparks-public-outcry/.
43.	https://www.statnews.com/2020/03/23/coronavirus-i-asked-eight-chatbots-whether-i-had-covid-19/.
44.	https://www.ft.com/content/0aafc2de-f46d-4646-acfd-4ed7a7f6feaa.
45.	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7186767/.
46.	https://web-assets.bcg.com/f1/79/cf4f7dce459686cfee20edf3117c/mit-bcg-expanding-ai-impact-with-organizational-learning-oct-2020.pdf.
47.	https://www.weforum.org/agenda/2020/09/ai-is-here-this-is-how-it-can-benefit-everyone/.
48.	http://www3.weforum.org/docs/WEF_Global_Technology_Governance.pdf.
49.	http://www3.weforum.org/docs/WEF_Agile_Regulation_for_the_Fourth_Industrial_Revolution_2020.pdf.
50.	https://www.weforum.org/reports/chatbots-reset-a-framework-for-governing-responsible-use-of-conversational-ai-in-healthcare.
51.	http://www3.weforum.org/docs/WEF_Responsible_Limits_on_Facial_Recognition_2020.pdf.
52.	https://www.weforum.org/reports/ai-procurement-in-a-box.
53.	https://www.weforum.org/projects/model-ai-governance-framework.
54.	http://www3.weforum.org/docs/WEF_Reimagining_Regulation_Age_AI_2020.pdf.
55.	https://www.weforum.org/projects/human-centred-artificial-intelligence-for-human-resources.
56.	https://www.nature.com/articles/s42256-021-00298-y.
57.	https://www.ipcc.ch/.
58.	https://www.sciencemag.org/news/2014/07/scientists-call-limit-creating-dangerous-pathogens
59.	https://law.stanford.edu/2015/04/04/of-science-crispr-cas9-and-asilomar/
60.	https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages.
61.	https://www.fhi.ox.ac.uk/windfallclause/.
62.	https://www.nationalgeographic.com/environment/article/artificial-intelligence-climate-change.
63.	https://www.nature.com/articles/s42256-020-0219-9?proof=t.

WORLD
ECONOMIC
FORUM

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.