

# AI ethics in action

An enterprise guide to progressing trustworthy AI



# How IBM can help

Clients can realize the potential of AI, analytics, and data using IBM's deep industry, functional, and technical expertise; enterprise-grade technology solutions; and science-based research innovations. For more information about AI services from IBM Consulting, visit ibm.com/services/artificial-intelligence

For more information about AI solutions from IBM Software, visit ibm.com/Watson

For more information about AI innovations from IBM Research®, visit research.ibm.com/artificial-intelligence



## Key takeaways

Organizations can distinguish themselves with a strategic approach to ethical issues in AI.

## Business leaders are taking accountability for AI ethics in today's enterprise

Non-technical executives are now the primary champions for AI ethics, growing from 15% in 2018 to 80% 3 years later—and 79% of CEOs are now prepared to act, up from 20%.

## Many organizations have made solid strides toward purposeful AI

More than half of organizations have taken steps to embed AI ethics into their existing approach to business ethics—and many of those are creating AI-specific ethics mechanisms.

# But the intention-action gap is still too wide

For example, having a diverse and inclusive workforce is important to mitigating bias in AI—acknowledged by 68% of organizations—but AI teams are still substantially less diverse than their organizations' workforces: 5.4 times less inclusive of women, 4 times less inclusive of LGBTQ+ individuals, and 1.7 times less racially inclusive.

# The intensifying imperative for advancing trustworthy AI

Business leaders need only scan the headlines to find examples of companies confronted with various societal, environmental, and political issues.

Customers, employees, and even shareholders are more frequently demanding that organizations not only take a principled stance on current concerns, but also follow through with meaningful actions that lead to clear outcomes. (See the Perspective "Companies that do good can do better, if they do it right.")

Executives today are increasingly impacted by these forces, many of which lie outside their usual comfort zones. Indeed, nearly 60% say recent societal and economic upheaval is expected to impact the broader social contract.<sup>1</sup>

These dynamics extend into the realm of AI and data, as the IBM Institute for Business Value (IBV) first explored in a survey conducted in 2018.<sup>2</sup>

For the general population, the definition of *un*trustworthy AI may be obvious: discriminatory, opaque, misused, and otherwise falling short of general expectations of trust. Yet advancing trustworthy AI can remain challenging considering the pragmatic balancing act sometimes needed: for example, between "explainability"—the ability to understand the rationale behind an AI algorithm's results—and "robustness"—an algorithm's accuracy in arriving at an outcome.

Organizations can no longer adopt AI without also addressing these tradeoffs and other ethical issues. The only question is whether they confront them strategically, purposefully, and thoughtfully—or not.

The broadly defined technology sector has taken the first steps. These businesses include many digital natives with high-growth, highly profitable platforms that can subsidize the ethical toll of their consumer-oriented business models. Consequently, they have assembled hefty teams and many-layered processes to address the demands, with varying degrees of success.

## **Perspective**

Companies that do good can do better, if they do it right

A company's traditional remit to boost *shareholder* value is increasingly too narrow. Broader demands from consumers, citizens, workers, and investors—amplified by the media—are causing organizations to pay more attention to *stakeholder* concerns that impact business results.

This growing trend is evident in the IBV's sustainability research: 2 years ago, only 8% of consumers were willing to pay more than a 100% premium to support brands with a purpose.<sup>3</sup> More recently, 43% said they are willing to pay that premium to support sustainable and environmentally responsible brands.<sup>4</sup>

Among employees, nearly 70% said they are more likely to accept a job offer from an organization they consider to be environmentally and socially responsible, and a similar dynamic impacts retention.<sup>5</sup>

Moreover, 68% of organizations' customers and 62% of their workers indicated they have the power to force corporations to change by increasingly voicing their views.<sup>6</sup>

Individual investors too are factoring sustainability into their financial decisions. Half said a company's climate change exposure affects its financial risk. And 92% of this group expected to invest, divest, or lobby fund managers to change investment mixes based on environmental factors and/or social responsibility in the next 12 months.<sup>7</sup>

Indeed, re-envisioning traditional shareholder capitalism to address moral and ethical issues important to a wider array of stakeholders is also gaining momentum.<sup>8</sup>

Yet, it's not as simple as adopting the latest buzzy mantras. The companies that are successfully realizing benefits from these market trends are viewing sustainability not as just another societal wave to ride—but as a catalyst for business transformation. Only then can they begin doing the hard work of embedding meaningful practices throughout the enterprise.

Done right, corporate values aren't just platitudes; they can lead to real financial value.

Incumbents from other industries must also navigate this uncertain environment, even while striving to drive profitable growth using AI in selected parts of their businesses. Adding to the difficulty, they often have fewer resources to dedicate to the ethical issues raised by AI.

Organizations face a stark choice.

They can batten down the hatches—hoping that the storm of consumer, employee, and investor engagement wanes, perhaps seeking shelter in the more familiar haven of regulation. Yet regional, national, and even local legal frameworks for governing AI continue to evolve. So, betting on compliance as a strategy in an uncertain (and overlapping) regulatory environment could be more of a gamble than it otherwise seems. The hazards of inaction may only continue to grow.

Alternatively, organizations can set a direction to distinguish themselves from competitors by taking a strategic approach to ethical issues in AI. And as they learn by doing, adjustments to their specific course of action are inevitable.

To understand how far organizations have progressed in their efforts to address and incorporate AI ethics into their business practices, we partnered with Oxford Economics in 2021 to survey 1,200 executives across 16 business and technology roles in 22 countries. (See "Study approach and methodology" on page 28.)

A meaningful approach to the discipline of AI ethics can yield the outcome of advancing trustworthy AI.

# A disquieting disparity between intent and implementation

To start at the beginning, let's define AI ethics.

AI ethics is generally recognized as a multidisciplinary field of study that aims to optimize the beneficial impact of AI by prioritizing human agency and well-being while reducing the risks of adverse outcomes to all stakeholders. A meaningful approach to the *discipline* of AI ethics can yield the *outcome* of advancing trustworthy AI.

Parallels exist in other professions. For example, the field of medicine is dedicated to improving patient health, while safe, effective, and trusted treatments are the means to achieving those ends. Engineering is a discipline focused on designing and constructing physical infrastructure, while safe and secure bridges, tunnels, and buildings are the intended result.

Other IBV research reveals that consumers, citizens, and workers view AI as a top technology to help solve the most pressing challenges facing humanity. 

Thus, a rigorous approach to AI ethics would seem essential. Furthermore, more than 85% of these respondents indicate that it is important for organizations to address AI ethics while tackling society's problems. 

2

Yet these same individuals also give companies in traditional industries, such as banking, retail, and insurance, mediocre marks for the responsible use of technology. Most of these industries also fall below the grades given to those social media platforms frequently maligned in the press (see Figure 1).<sup>13</sup> Moreover, there has been little apparent improvement since 2018, when 40% of consumers trusted companies to be responsible and ethical in developing and implementing new technologies such as AI —a similar average across industries today.<sup>14</sup>

More than 85% of surveyed consumers, citizens, and employees indicate that it is important for organizations to address AI ethics.

#### FIGURE 1

## AI ethics report card

Consumers gave mediocre grades to traditional industries in the responsible use of technology

## Percent earning a grade of A or B

. 0.00.	in carning a grade of A of B		
56%	Hardware/software firms	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	Consumers rate <b>technology</b>
51%	Streaming entertainment platforms	503	<b>companies</b> highest
47%	Telecommunications		8 - 1
46%	Automotive manufacturers		
44%	Healthcare providers		Even social media
43%	Social media platforms		firms rank higher than many traditional industries
41%	Banking & financial services providers	•	
41%	Educational institutions		
39%	Travel & transportation providers		
38%	Traditional media & news outlets		
38%	Utility companies		
35%	Insurers		
34%	Retail merchants		
34%	Government organizations		

Q: What grade/rating would you give the following types of companies/organizations for their responsible use of technology like artificial intelligence?

Source: IBM Institute for Business Value Human Insights Global Survey of 14,526 adults. July 2021. Previously unpublished data.

However, executives surveyed for this study say their organizations *do* view AI ethics as more important now than 3 years ago—increasing from less than half the respondents in our 2018 survey<sup>15</sup> to more than three-quarters in 2021. More concretely, they estimate their spend on training, teams, processes, tools, and other operational capabilities to institutionalize AI ethics has doubled over that period. And they forecast greater investment over the next 3 years.

So, how to explain the disparity between these executives' views and the assessments of their customers and employees?

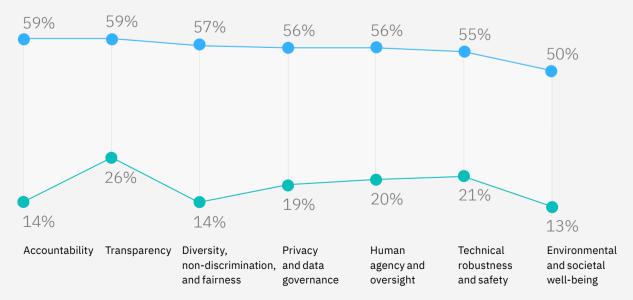
Organizations' aspirations—intensified by the urgency to meet market expectations—seem to be exceeding their ability to put intent into practice quickly enough to appease stakeholders. While more than half of organizations have publicly endorsed common principles of AI ethics, less than a quarter have operationalized them (see Figure 2). Fewer than 20% strongly agree that their organizations' practices and actions on AI ethics match (or exceed) their stated principles and values. This confirms and quantifies what the World Economic Forum (WEF) calls the "intention-action" gap.<sup>16</sup>

Fewer than 20% of executives strongly agree that their AI ethics actions meet or exceed their stated principles and values.

### FIGURE 2

## The intention-action gap

Organizations are endorsing AI ethics principles—but are still catching up on implementing them



## Endorsed | Operationalized

Note: AI ethics principles as defined by the European Commission High-Level Expert Group on AI in "Ethics guidelines for trustworthy AI." April 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Closing these operational gaps is critical. For example, for AI to become less biased and more trustworthy, the ethics principle of diversity, non-discrimination, and fairness must be addressed. Organizations recognize the importance, rating this item as significantly more important to their AI efforts in 2021 than in 2018. Virtually no leaders rated this as unimportant in the latest survey, while almost 6 times more executives rated it "very important."

To address this issue, employees working on AI ethics should represent the broader audience; however, AI teams remain substantially less diverse than the organizations' total workforces. Our results reveal 5.4 times fewer women on the AI teams than in the organization, along with 4 times fewer LGBTQ+ individuals, and 1.7 times fewer Black, Indigenous, and People of Color (BIPOC) (see Figure 3).

Root causes for these differences vary. However, a more proactive and tangible focus on improving recruiting, hiring, and retention practices is needed to increase underrepresented voices among those designing and implementing AI. This emphasis can help mitigate bias and build trust in AI outcomes.

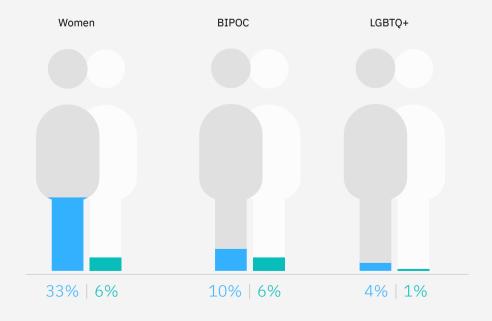
We see some evidence in our survey that these efforts pay off financially too. For example, businesses in the quartile with the smallest gender gap between their workforces and their AI teams realize a somewhat higher return on investment for their AI projects.

Similar efforts with tangible goals and measurable results are required to remove disparities between other AI ethics principles and practices. (See the case study "Regions Bank: A focus on high quality and trustworthy AI.")

### FIGURE 3

## **Diversity dichotomy**

Organizations' AI teams are significantly less diverse than their enterprise workforces



Enterprise | AI

The discrepancy between aspirations and actions parallels companies' approaches on environmental sustainability. In recent IBV research, only 35% of organizations have acted on their sustainability strategy, and only 37% have aligned sustainability objectives with their business strategy. Just one-third have integrated sustainability objectives and metrics into business processes.

When organizations and executives focus more on putting AI ethics and other principled priorities into practice, they can more successfully address the intention-action gap.

Executives cite 20 different business functions as at least somewhat involved in AI ethics, suggesting synergy is essential.

#### FIGURE 4

#### Changing of the guard

From 2018 to 2021, those primarily accountable for AI ethics have shifted from technical to non-technical leaders

## 2018 | 2021

#### Non-technical leaders Technical leaders +24 pts +5 pts +9 pts -27 pts -33 pts +5 pts -2 pts -11 pts 34% 31% 28% 14% 10% 10% 8%\* 7%\* 6% 5% 5% 4% 3% 3% 1% 1% CEO CISO Board General Privacy Risk & Data CIO СТО CDO Officer Counsel Compliance Science/ Officer AI/ML

Q: Which function is primarily accountable for AI ethics? Source for 2018 survey data: Goehring, Brian, Francesca Rossi, and Dave Zaharchuk. "Advancing

AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020.

## A seismic shift: Business executives lead, and collaboration is vital

As organizations work toward closing the gap, opinions on who is responsible for spearheading that effort have shifted dramatically in the last few years. In 2018, only 15% of respondents pointed to a non-technical executive as the primary "champion" for AI ethics; 19 in 2021, 80% do.

Companies are looking especially to CEOs (28%)—but also to Board members (10%), General Counsels (10%), Privacy Officers (8%), and Risk and Compliance Officers (6%) to lead the way—a finding that varied little by industry or geography (see Figure 4). With this new expectation, CEOs may want to consider appointing a dedicated AI ethics executive to share the responsibility.<sup>20</sup>

<sup>\*</sup>Position was not included in 2018 data

#### Case study

Regions Bank: A focus on high quality and trustworthy AI<sup>21</sup>

How can an enterprise operationalize AI in a way that is repeatable, sustainable, and, perhaps most important, trusted?

Regions Bank had to face this challenge. Its advanced analytics practice too often relied on siloed data sets, development teams working in isolation, and disparate, somewhat inconsistent development methods.

First, the bank transformed its analytics function. After creating an analytics center of excellence, it brought data into a centralized environment and applied more machine learning (ML) and AI techniques. Above all, it adopted an end-to-end business value approach that includes AI quality control.

With this new foundation, Regions Bank can better leverage an integrated set of capabilities and teams to help ensure that AI models are fair, ethical, and trusted.

"Regions prides itself on open and trust-based interactions with our customers," writes Manav Misra, Chief Data and Analytics Officer. Regions Bank wants to improve the lives of its communities, customers, and associates, Misra explains. "That's why we bake AI ethics oversight into our development methodology."

He adds, "All this points to what we call 'Responsible AI.'" This approach requires the data underlying AI models to be representative of the data used to make decisions. Plus, the models must be explainable, so the decision-making process is clear.

An internal oversight team helps uphold the fairness, safety, and soundness of the solutions, as do risk managers, audit partners, and government regulators.

But when is the right time to engage AI ethics processes and teams?

The sooner the better, according to Misra. "If we can get our oversight partners onboard early to understand the business case and requirements, along with the code and the developers' intent throughout every agile sprint, they can provide faster feedback," he writes. "This accelerates development of the high-quality, trustworthy AI solutions that Regions Bank strives for."

The result has been more highly trusted AI, machine learning, and other analytics solutions. Together, they help reduce risk, help detect fraud, assist commercial customers, and provide insights into consumers to better meet their needs—all of which delivers business value.

That said, organizations still recognize a crossfunctional, collaborative approach is essential. Most leaders name at least 20 different business functions that are somewhat involved in AI ethics.

Chief Information Officers (CIOs), Chief Technology Officers (CTOs), Chief Data Officers (CDOs), and their teams are essential to operationalizing AI ethics. But these duties do not fall only in the realm of technical experts. Additional functions that are important in AI ethics include procurement, product design, research, public policy, and regulatory affairs.

"Organizations need intense collaboration to make AI ethics mechanisms real," says Rob Reich from Stanford University's Institute for Human-Centered Artificial Intelligence (HAI), in response to these findings. "Just as people can't outsource personal ethical decisions to others, organizations can't pursue AI ethics by installing a Chief Ethics Officer and having all other business units continue as usual. At some basic level, ethics has to be everyone's responsibility."<sup>22</sup>

By way of example, a large global pharmaceutical company initiated a modest effort within its data science team to develop a proof-of-concept skills inference model that could help inventory and predict talent development needs. Based on initially positive results, the company decided to scale the project across the enterprise. Yet it quickly became apparent that achieving AI at scale demanded far more participation than data scientists.

"At some basic level, ethics has to be everyone's responsibility."

**Rob Reich,** Associate Director, Institute for Human-Centered Artificial Intelligence, Stanford University It required a collaborative effort among HR professionals, legal teams, psychology experts, and other talent management professionals—in addition to software engineers, project managers, and other IT specialists. And in this case, data lineage and provenance were not just optional metadata attributes; they were crucial elements for building trust in the outcomes, as well as driving employee adoption and engagement.

In this context—where involvement from multiple business functions, a heightened level of sensitivity, and dynamic feedback loops are prevalent—top-down sponsorship is critical to facilitating the right interactions among the right teams at the right times.

Indeed, the executives at the top seem better prepared than in 2018. Perhaps emboldened by engagement on the topic with their teams, outside experts, and in conference settings, a fourfold increase in CEOs (from 20% in 2018<sup>23</sup> to 79% in 2021) and more Board members (from 45% in 2018<sup>24</sup> to 71% in 2021) say they are prepared to act on AI ethics issues.

Chief Human Resources Officers (CHROs) also appear to be doing a better job addressing a deficit we identified in 2018. At that time, only 37% of organizations had active plans to retrain and reskill workers impacted by AI.<sup>25</sup> In 2021, this number jumped to 55%. Similarly, organizations with learning plans for employees who need to interact more with AI increased from 41% in 2018<sup>26</sup> to 70%. As the population of those who need to engage directly with AI rises—estimated by CHROs to grow by over 20% in the next few years—proactively addressing these education issues is increasingly important for companies, governments, and societies.

The anticipation of external AI governance regulations also may be elevating the importance of AI ethics. 74% of organizations report that regulations are imminent in their region—and 64% say they are prepared to address them. However, independent of the way the regulatory environment unfolds, respondents indicate that third-party professional bodies and other non-regulatory groups are expected to play a key role, with 48% citing the influence and involvement of these broader ecosystems as very important. (See the case study "A Canadian banking consortium: Jointly building trusted data and AI.")

Overall, organizational engagement trends appear to be headed in the right direction. This provides momentum for companies as well as government and educational institutions to *do* more to implement AI ethics principles.

The number of organizations with active plans to retrain and reskill workers impacted by AI jumped from 37% in 2018 to 55% in 2021.



## Case study

A Canadian banking consortium:
Jointly building trusted data and AI<sup>27</sup>

Financial services institutions are using data—including customer, internal, and third-party data—to deliver innovative alternatives to traditional banking solutions. As a result, the bank-customer relationship could become much deeper and more interactive.

As Mathieu Avon, Vice President, Integrated Risk Management, National Bank of Canada, said, "AI can positively transform the client and employee experience and augment the core capabilities of organizations. We also have a key role to play in deploying AI technologies responsibly and ethically in order to maintain the trust of clients and other stakeholders."

Many financial services organizations agree that ethically aligned AI is a necessity—in principle. But they don't always have the resources to build, execute, and scale trusted data and AI applications on their own.

To respond to that challenge, an ecosystem of partners—convened by the Institute of Electrical and Electronics Engineers (IEEE), the world's largest technical professional organization—joined forces in 2019 to create a standards guide focused on ethically aligned AI for financial services in Canada.

"Data ethics aren't a cost of doing business; they are an investment in good business," said Terry Hickey, former Senior Vice President and Chief Information Officer, Enterprise Data, CIBC.

The team curated best practices and roadmaps from industry, academia, nongovernmental organizations, and standards bodies. 6 major banks in Canada contributed to the initiative, as did credit unions, pension funds, and fintechs.

William Stewart, Head of Data Use and Product Management, Data and Analytics, Royal Bank of Canada, highlighted the link to corporate principles: "Organizations will need to demonstrate that they have thoughtfully considered the unique risks of AI and responded in a manner that is consistent with their corporate values and aligned with the expectations of clients, employees, and society."

This ecosystem has continued to collaborate, creating a playbook for trustworthy machine learning operations processes and tools. The shared outcome from this ongoing effort is intended to reassure customers, partners, and employees that privacy is protected, data is used responsibly, and bias is mitigated.

# Impetus to action: Trustworthy AI can deliver business value

The business justifications for moving from principles to practice are not just about mitigating downside risks and meeting stakeholders' growing expectations for acting responsibly. Companies can potentially realize additional benefits. Achieving sustainable practices and guarding information quality—as well as embracing long-term thinking over short-term profits—tangibly increases trust in consumers' minds.<sup>28</sup>

Likewise, we found in our research that 75% of executives view ethics as a source of competitive differentiation. A similar percentage translates that advantage into pragmatic decision criteria for business partner and vendor selection, including in the realm of AI. More than half are willing to pay a price premium to companies they deem ethical.

An organization's positioning relative to its peers on this issue appears to be motivated primarily by 2 factors: (1) the importance of AI, and (2) the importance of ethics in AI.

We found organizations that view AI as important to their business strategy are 1.5 times more effective with their AI initiatives, based on their assessments. They also achieve 2 times the ROI on their AI projects relative to those companies that view AI as less important. The strategic importance of AI corresponds to greater company-wide revenue growth and profitability as well.

Moreover, those organizations that place greater importance on AI ethics reported that they also have a greater degree of trust from customers and employees.

So strategic intent—supported by pragmatic action—may help an enterprise realize greater business value from AI.

But how are organizations responding?

75% of executives view ethics as a source of competitive differentiation.

Companies that view AI as important to their business strategy report that they achieve 2x ROI on their AI projects relative to those that view it as less important.

Figure 5 shows that those currently viewing AI and AI ethics as less important expect to move over the next 3 years to the upper right quadrant, where the most advanced adopters are driving leading practices and achieving greater impact. The number of organizations in this segment is expected to double in that timeframe, based on their executives' stated intent.

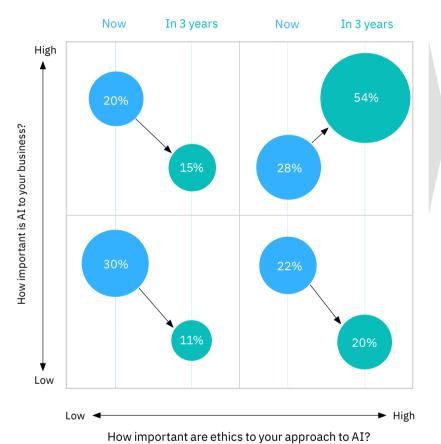
We view the upper left quadrant—a position where AI is important, but AI ethics are less so—as unsustainable, especially given the consumer, citizen, investor, and employee dynamics identified earlier in this report. Organizations cannot afford to remain uncommitted to AI ethics. Just as social media has enabled a world where teenage foibles are recorded for all time, businesses are creating historical records of their behavior to be evaluated by partners and prospective employees—now and into the future. Concerted efforts (and investments in capabilities) can help organizations move to a more viable position.

This research is another example of where improved financial performance correlates with strategic and effective use of AI—as many previous studies have shown.<sup>29</sup> Now it appears the relationship also applies in the context of thoughtful, robust approaches to adopting AI ethically. That may be why some companies with AI and AI ethics core to their business strategies are striving to embed practices deeply throughout their organizations. (See the Perspective "Leading practices from the field" on page 16.)

## FIGURE 5

## **Positioned for progress**

The number of organizations embracing the value of AI and AI ethics is expected to grow over the next 3 years



In the next 3 years, a majority of companies expect both AI and AI ethics to be very important strategically

### **Perspective**

Leading practices from the field: The Responsible Use of Technology project

The World Economic Forum (WEF), along with the Markkula Center for Applied Ethics at Santa Clara University, has introduced the Responsible Use of Technology case studies initiative. With this project, the WEF enables organizations to learn from the insights of its community and take steps to incorporate ethics into their design, development, and use of technology.

The first 2 studies focus on Microsoft and IBM, exploring the companies' evolving efforts to cultivate ethical values and culture. The study on Microsoft covers the tools and practices in its product engineering organization.<sup>30</sup> The white paper on IBM examines how it has transformed to think about technology ethics more deeply and to design, develop, deploy, and use technology more responsibly.<sup>31</sup>

With AI ethics, Microsoft focuses specifically on its R&D capabilities, while IBM expands more broadly across the enterprise. Both take the following AI-related actions:

- Endorse principles aligned with the companies' values and those of outside multi-stakeholder groups
- Operationalize through formally published practices and standards, design thinking techniques, and other pragmatic toolkits for practitioners
- Measure with goal setting and performance evaluations
- Enable internally with ambassadors throughout the business and through training programs—and organically with a strong emphasis on culture
- Augment externally by partnering with outside organizations and participating in cross-industry, government, and scientific initiatives
- Sustain with deep research in the field and continuous improvement, grounded in the impact on society.

Microsoft's AI governance approach follows a hub-and-spoke model that helps the company integrate privacy, security, and accessibility features into its products and services. The hub consists of 3 teams: the first comprising scientific and engineering experts; the second focusing on policy, governance, enablement, and sensitive use; and the third enabling responsible use processes among the engineering teams. Meanwhile, the spokes consist of business experts, including a design thinking team and "champs" that help drive cultural shifts.

The IBM AI Ethics Board serves as the keystone for the ethical development and deployment of AI systems. The Board, which includes senior leaders from the business units and corporate functions, has the authority to implement, enable, and enforce its decisions. It also has a project office and an advocacy network that jointly drive change throughout the organization through local focal points.

This governance approach includes IBM-developed AI and supporting software, business partner relationships, client deployments assisted by its services organizations, and internal use (such as hiring, compensation, education, and diversity and inclusion efforts).

Both companies have taken steps to be responsive when they discover inconsistencies between their values and outcomes. They also continue to evolve their approaches in the spirit of dynamic feedback loops and continuous improvement.

Organizations in other industries might counter that their needs differ from the structures and mechanisms put in place by large technology companies. While that's sometimes true, there are leading practices, lessons learned, and other success factors to consider and perhaps scale down while preserving much of their robustness.

As the WEF asserts, "Because all companies are now technology companies, all companies should think more closely about how technology ethics is involved in their work." 32

Companies that view AI as central to their business models have embraced governance approaches specific to AI ethics. Organizations where AI features less prominently in their strategy—but is still being deployed in non-core functions or less central business units—can get started by integrating AI ethics principles into existing business ethics mechanisms (see Figure 6).

Correlations between the inclusion of AI in the business strategy and the effectiveness of AI are not necessarily causal, of course, but organizations should be wary of steering away from where the market is headed.

#### FIGURE 6

#### First steps

Many organizations are incorporating AI ethics into existing business ethics mechanisms

57%	Business conduct guidelines				
49%	Periodic mandatory training and educational materials to refresh and reinforce policies				
48%	Risk assessment framework and auditing/ review process for high-risk projects				
47%	A mission/values statement that is clearly communicated to all employees				
46%	Buying criteria/due diligence for vendor engagement				
46%	Anonymous employee hotline				
46%	An actively supported culture of ethical decision-making				
. 676	An actively supported culture of ethical				
46%	An actively supported culture of ethical decision-making  Tools and other materials to support ethics				
46%	An actively supported culture of ethical decision-making  Tools and other materials to support ethics diagnostics and decision-making				

# Meet the moment: Applying academic ideals to real-world actions

As organizations launch their AI ethics initiatives, they can benefit from an abundance of resources. The field of AI ethics—initially explored in cross-disciplinary research settings—offers many constructive frameworks, assets, and associations to leverage (see Figure 7).

## FIGURE 7

## Framing the future

The field of AI ethics enters a more pragmatic phase

	0.0	1.0	2.0	3.0
Focus	Awareness	Principles	Policies	Practice
Timeframe	Up to 2016	2017-19	2020-21	2022 and beyond
Approach	Small, multi-disciplinary academic groups	Multi-stakeholder involvement in AI research community	Inter-organization engagement across AI developers/companies	Cross-/intra-organization, society-wide engagement
Communities	AI/computer science, philosophy, law, economics	AI/tech firms, think tanks, standards bodies, AI developer community	Corporate boards, professional bodies, policymakers, regulators	Business executives, functional users, customers, accreditation groups
Working group examples	- Puerto Rico conference on Future of Life	<ul> <li>AI research community</li> <li>Asilomar Conference on Beneficial AI</li> <li>European Commission (EC) High-Level Expert Group on AI</li> <li>Partnership on AI</li> <li>Organisation for Economic Co-operation and Development (OECD) AI Principles</li> </ul>	<ul> <li>WEF Global AI Alliance</li> <li>US National Institute of Standards and Technology AI end user trust initiative</li> <li>Global Partnership on AI</li> </ul>	<ul> <li>Companies initiating cross-functional teams</li> <li>Governments finalizing regulations</li> <li>Standards defined and released</li> </ul>
Example outputs	<ul> <li>Academic papers</li> <li>Conference presentations</li> </ul>	<ul> <li>Organizations' stated principles</li> <li>WEF Board Toolkit</li> <li>EC Ethics Guidelines for Trustworthy AI</li> <li>US General Services Administration Transformation Technology Services handbook</li> </ul>	<ul> <li>Organizations' stated policies</li> <li>IEEE Trusted Data &amp; AIS Playbook</li> <li>EC White Paper</li> <li>Personal Information Protection Law (China)</li> <li>New Generation AI Ethics Specification (China)</li> </ul>	<ul> <li>Organizations' ethics in action</li> <li>Company training</li> <li>Professional certification</li> <li>EU AI regulation (proposed)</li> <li>China AI regulation (proposed)</li> </ul>

Even university dons acknowledge that their objective has always been for the academic research and abstract concepts to find their way into pragmatic use—and not just in the realm of computer science. "AI ethics can't be done from an armchair," acknowledges Stephen Cave, the Director of the Leverhulme Centre for the Future of Intelligence and a Senior Research Associate in the Faculty of Philosophy at the University of Cambridge. 33 His university is piloting a bespoke graduate degree beginning in 2021 to help foster the next generation of AI ethicists.

Several universities have created institutes with a similar focus (for example, the Institute for Ethics in AI at Oxford University). Related courses for students in other educational settings are increasing as well. Think tanks such as the Brookings Institute and the Future of Privacy Forum; professional organizations such as IEEE; non-regulatory governmental bodies such as the US National Institute of Standards and Technology; and cross-industry groups such as the WEF, the Global Partnership on Artificial Intelligence. and the Partnership on AI also have been at the forefront of AI ethics-and have created useful resources. Other grassroots collaborative efforts (for example, the US Global Services Administration's Technology Transformation Services) have emerged along the way.

More than 67% of organizations that view AI and AI ethics as important also indicate they outperform their peers in sustainability, social responsibility, and diversity and inclusion.

Organizational leaders can now leverage these advancements to help operationalize the discipline of AI ethics across the many enterprise functions that need to work together. Ultimately, these tools can help advance the delivery of trustworthy AI.

This approach to AI ethics also can complement organizations' efforts to achieve competitive advantage, sustainable innovation, and even social justice goals through the responsible use of technology. Indeed, our research suggests a relevant connection: more than 67% of organizations that view AI and AI ethics as important also indicate they outperform their peers in sustainability, social responsibility, and diversity and inclusion.

"Applying algorithms with great responsibility is paramount for Dutch government agencies," notes Johan Maas, CIO of the Netherland Enterprise Agency RVO. "By using a holistic approach, our AI not only advises employees about our clients' subsidy applications, but also explains those parameters that were important to its advice—which lays the foundation to create an effective human-AI collaboration."

He continues, "This is in line with our policy of developing AI according to principles like equal treatment, which conforms to ethical and democratic principles and transparency. Accordingly, we increase our employees' and citizens' trust in using AI."<sup>34</sup>

The stage is set. More businesses are prepared to embrace a stakeholder-first mindset. Consumers continue to say they will pay premiums for products viewed as sustainable. Many procurement functions are increasingly integrating criteria for business

responsibility into sourcing decisions. Corporate ethics and purpose are viewed as differentiators. And many people are acknowledging the growing impact of AI on the social contract. Building trustworthy AI through a formal, integrated approach to AI ethics must be more than a noble ambition.

It is a strategic and societal imperative.

And now is the time for organizations to do more.



## Action guide

Next steps depend on where your organization is in adopting AI and your corresponding strategic and ethical intents.

For those who do not yet view AI as essential to the business, keeping a close eye on competitive shifts (and the regulatory environment) is crucial to evaluating how to adopt AI more proactively, and ethically. Leveraging existing business ethics mechanisms shown in Figure 6 on page 18 can be a reasonable starting point.

For organizations further along on their AI journey and with an ambitious future agenda, an AI Ethics Board that governs enterprise-wide AI can be a natural catalyst to adopting the practices necessary to help build trustworthy AI. And embedding ethics throughout the AI lifecycle—from design to deployment to dynamic feedback—is essential to operationalizing those practices.

For those already approaching AI and AI ethics strategically, expanding your approach to include suppliers, customers, and other ecosystem partners can enable a greater degree of "ethical interoperability"—a term coined by David Danks, a Data Science & Philosophy Professor at the University of California San Diego.<sup>35</sup> This could help you differentiate further from your competitors and protect the business from backdoor risks.

The next steps also depend on who you are.
CEOs and other C-suite executives help set the direction, business teams make decisions based on AI-generated insights, data scientists interrogate the data, AI developers construct the solutions—and many others fill additional positions.

Just as AI is a team sport, so too is the field of AI ethics, but with a more diverse set of teams, a larger arena, and a heightened spirit of collaboration. And it's not just a technological challenge, but a sociotechnological challenge, so a holistic approach is necessary. <sup>36</sup> Begin by thinking through the following 3 high-level focus areas, and then review the role-specific recommended actions.

## Organization-wide actions

## 01

## Strategy and vision

## Set ethical AI practices in the proper strategic context

- Consider the criticality of building trustworthy AI to business strategy and objectives:
  - What are key value creators that could be accelerated with AI?
  - How will success be measured?
- Consider the role of AI innovation in an organization's growth strategy and approach (for example, organizations considered trailblazers, fast followers, and so on).
- Endorse key principles of AI ethics.
- Determine human + machine balance in the organization.<sup>37</sup>

Key roles: CEO, business and technology executives, Board of Directors

## $\odot 2$

## Direction and management

## Establish a governance approach for ethical AI implementation

- Incorporate stakeholder perspectives (for example, leaders, employees, customers, government).
  - Address dimensions of privacy, robustness, fairness, explainability, transparency, and other relevant principles important to your organization.
  - Consider the enterprise and broader ecosystem.
- Set an AI and data risk profile and threshold level.
- Establish an organizational structure, policies, processes, and monitoring.

Key roles: AI Ethics "Champion," risk and compliance executives and leaders, legal counsel, HR/talent executives and leaders, diversity and inclusion executives and leaders, business unit/functional executives and leaders, and technical leaders

## 03

## Implementation and deployment

## Integrate ethics into the AI lifecycle

- Engage with key stakeholders (for example, leaders, employees, customers, government).
- Establish organizational structure, policies, processes, and monitoring.
- Capture, report, and review compliance data.
- Drive and support education and diversity efforts for teams.
- Define integrated methodologies and toolkits.

Key roles: AI ethics program office, business and technology leaders and teams, HR/talent leaders and teams, AI designers, AI engineers, data scientists, and other technical teams

# Role-specific actions

## Key roles **Example actions** CEO - Marshal an executive team around relevant strategic and business objectives to drive the innovation strategy for AI. - Allocate capital appropriately to support AI objectives. - Convey your organization's strategic point of view, values, and corresponding actions on business and AI ethics through blogs, keynotes, internal communications, and other appropriate tools. - Consider appointing a Chief AI Ethics Officer to lead enterprise-wide efforts, depending on the strategic importance of AI, and/or make accountability clear among current executive roles.38 - Consider establishing an AI ethics board, depending on the strategic importance of AI (and chair or partner alongside the board chair). - Report progress to the Board of Directors and other key stakeholders. - Approve an implementation approach for AI ethics. - Promote the training, AI and data literacy, and change management agenda. - Sponsor the overall AI and data risk profile and update cadence. - Align executives to common AI ethics goals across business units and functions. **Business and** - Prioritize areas for AI, machine learning, analytics, and data in the context of your technology enterprise-wide innovation strategy to optimize business impact. executive teams - Affirm values as part of your corporate identity and culture. - Design the approach to human + technology collaboration. - Define macro success metrics and the timeframe for achieving them. - Augment your innovation ecosystem by identifying and engaging key AI-focused technology partners, academics, startups, and other business partners to establish "ethical interoperability." **Board of Directors** - Ratify the strategic direction. - Hold the executive team accountable through engagement with the full Board and key committees (for example, Risk/Audit Committee, Technology Committee). **AI Ethics** - Act as guide and coach. "Champion" - Set the strategy, direction, and approach for the AI ethics program: and program office • Derive ethical implications for AI from organizational values. • Determine the operational and organizational approach (for example, a centralized Center of Excellence/Program Management Office, decentralized cross-functional/ cross-business unit advocates, funding, internal versus external resourcing, and so on). • Curate standards for AI models, such as user consent and opt-out implications. • Establish a regular process to review and audit procured and developed AI models. Assess the business, technology, and diversity balance in governance structures and consider rotating leadership and committee mix.

 Provide transparency by communicating incentives for engaging AI ethics mechanisms, escalations, whistleblower protections, and conflicts of interest

• Deploy AI ethics deeply into your organization's approach to AI (for example, create a playbook that details how to build a culture of responsibility).

Work with ecosystem partners to extend depth and reach.<sup>39</sup>

mitigation.

• Educate employees on how to mitigate risk.

24

## Key roles **Example actions** Risk and - Lead and coordinate an AI risk assessment and recommend program enhancements: compliance · Review existing governance frameworks and documentation, processes, and executives and controls for relevant programs, such as risk, compliance, privacy, security, records leaders management, and data governance. • Assess against new and emerging AI and data ethics risks in partnership with trained practitioners—balanced with a culture of open innovation. • Propose enhancements to address new and emerging AI and data ethics risks (for example, new committees, charters, processes, tools, mitigation efforts, risk insurance, and so on). · Lead and coordinate the build-out of new AI risk assessment programs, including roles, responsibilities, controls, escalation protocols, metrics, AI model audits, procedures, processes, and documentation. - Track global AI and data ethics developments, such as legal, regulatory, and policy changes: • Deliver regular summaries of changes and trend analyses. Build or integrate regulatory change management routines. **Chief Information** - Regardless of the degree of AI adoption, develop and augment security processes to accommodate AI-specific threats: **Security Officer** (CISO) · Identify and manage threats. • Document and provide compliance assurance. Legal counsel - Work with your partner and vendor ecosystem to examine the full chain of responsibilities and liability for AI applications and data uses. HR/talent - Develop a mitigation approach for identified functional risks. executives and - Collaborate to define the training, AI and data literacy, and change management leaders agenda. - Develop action plans for closing skill gaps and addressing impacted employee needs. Diversity and - Measure diversity on AI teams. inclusion - Collaborate to identify ways to stem attrition and to attract and retain executives and underrepresented groups in AI and other technology roles. leaders - Speak publicly on efforts to promote diversity in AI, data science, and related teams. - Establish training plans to help employees understand how unconscious bias can infiltrate data sets (and AI solutions). **Business unit** - Collaborate to define business unit- and function-specific implementation processes and functional and policies. executives and - Publicly and actively support ongoing AI ethics efforts across the organization. leaders

#### Key roles

#### Example actions

# Business and technology leaders and teams

- Orchestrate day-to-day implementation:
  - Create and augment AI teams and processes.
  - Determine roles for other employees and partners.
  - Identify outside experts needed, such as designers, ethicists, psychologists, and social scientists.
- Deploy end-to-end use of ethical AI practices:
  - Build ethical principles into business and technical requirements.
  - Assess challenges and opportunities with feedback from teams and work together to address.
  - Give teams time to focus on growing and improving ethical practices.
  - Align business value with key objectives and make them measurable for your teams; then set realistic goals to achieve.
  - Regularly communicate the benefits of working across disciplines to create ethical AI practices.
  - Cultivate organic "bottom-up" ideas to complement "top-down" approaches.
- Empower teams to raise and address issues if ethical principles are not met.

## HR/talent leaders and teams

- Curate and deliver tailored learning for AI ethics across the employee base.
- Conduct training for employees who need to work more with AI.
- Implement reskilling and upskilling plans for those impacted by AI.

Designers, architects, data scientists, social scientists, AI engineers, and other technical teams

- Leverage methods and tools to explore potential ethical concerns:
  - Lead design thinking sessions to identify and explore possible issues with the business and other experts well before code is written.<sup>40</sup>
  - Conduct exploratory data analysis on training data.
  - Set fairness thresholds, measure fairness metrics, and mitigate bias.
  - Communicate and demonstrate technical feasibility of "black box" transparency.
  - · Build interfaces to access algorithm explainability.
  - Create documentation to enable others to evaluate context.
  - Encourage teams to raise questions and address ethical concerns.
- Explicitly define components and functional requirements to define business success.
- Explain the need for AI ethics to non-technical audiences.
- Self-initiate critical training when not provided (for example, fairness and accuracy standards, data lineage and provenance mechanisms, adverse impact analysis).
- Conduct AI impact and risk assessments.
- Identify and communicate business and user stories that illustrate the importance of AI ethics.
- Share knowledge and collaborate across multiple teams (and types of teams).

## About

## the authors



## **Brian Goehring**

Global Research Lead, AI
IBM Institute for Business Value
goehring@us.ibm.com
linkedin.com/in/brian-c-goehring-9b5a453

Brian Goehring is an Associate Partner in the IBM Institute for Business Value, where he leads the AI business research agenda, collaborating with academics, clients, and other experts to develop data-driven thought leadership. He brings over 20 years' experience in strategy consulting with senior-level clients across most industries and business functions. He received an A.B. in Philosophy from Princeton University with certificates in Cognitive Studies and German.

#### Francesca Rossi

IBM Fellow, AI Ethics Global Leader Francesca.Rossi2@ibm.com linkedin.com/in/francesca-rossi-34b8b95 @frossi\_t Francesca Rossi is an IBM Fellow at the T.J. Watson Research Center and is an IBM AI Ethics Global Leader. In this role, she leads research projects to advance AI capabilities and co-chairs the IBM AI Ethics Board. Francesca is a member of the Board of Directors for the Partnership on AI and on the Steering Committee for the Global Partnership on AI. She is a Fellow for both the worldwide Association for the Advancement of Artificial Intelligence (AAAI) and the European Association for Artificial Intelligence (EurAI), and she was elected as the next president of AAAI. Before joining IBM, she was a professor of computer science at the University of Padova, Italy.

#### Beth Rudden

IBM Distinguished Engineer,
Cognitive Science and Trusted AI,
Data & Technology Transformation,
IBM Consulting
brudden@us.ibm.com
linkedin.com/in/brudden

Beth Rudden leads a global practice comprising large, geographically dispersed teams that develop AI models built on the ethical use of data to deliver actionable insights for clients. With a B.A. in Classical Archeology and an M.A. in Anthropology, she also holds several patents for solutions that provide clearer insights, better customer understanding, and faster implementation.

# IBM Institute for Business Value

For 2 decades, the IBM Institute for Business Value has served as the thought leadership think tank for IBM. What inspires us is producing research-backed and technology-informed strategic insights that help business leaders make smarter business decisions.

From our unique position at the intersection of business, technology, and society, we survey, interview, and engage with thousands of executives, consumers, and experts each year, synthesizing their perspectives into credible, inspiring, and actionable insights.

To stay connected and informed, sign up to receive IBV's email newsletter at ibm.com/ibv. You can also follow @IBMIBV on Twitter or find us on LinkedIn at linkedin.com/showcase/ibm-institute-for-business-value.

## About Research Insights

Research Insights are fact-based strategic insights for business executives on critical public and private sector issues. They are based on findings from analysis of our own primary research studies. For more information, contact the IBM Institute for Business Value at iibv@us.ibm.com.

## Related reports

## Advancing AI ethics beyond compliance

Goehring, Brian, Francesca Rossi, and David Zaharchuk. "Advancing AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020. https://ibm.co/ai-ethics

#### The business value of AI

Goehring, Brian and Anthony Marshall. "The business value of AI." IBM Institute for Business Value.

November 2020. https://ibm.co/ai-value-pandemic

#### Sustainability as a transformation catalyst

Balta, Wayne, Manish Chawla, Jacob Karl Dencik, and Spencer Lin. "Sustainability as a transformation catalyst: Trailblazers turn aspiration into action." IBM Institute for Business Value. January 2022. https://ibm.co/sustainability-transformation

## Study approach and methodology

In partnership with Oxford Economics, the IBM Institute for Business Value surveyed 1,200 executives in 22 countries across North America, Latin America, Europe, Middle East and Africa, and Asia (including China and India) from May through July 2021. The survey scope included 16 business and technology roles—primarily executives but also AI professionals—from more than 22 industries. The median revenue/budget for surveyed organizations was just under \$3B.

# Notes and sources

- Goehring, Brian and Anthony Marshall. "The business value of AI." IBM Institute for Business Value. November 2020. Previously unpublished analysis. https://ibm.co/ai-value-pandemic
- 2 Goehring, Brian, Francesca Rossi, and David Zaharchuk. "Advancing AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020. https://ibm.co/ai-ethics
- 3 Gupta, Sachin, Sashank Yaragudipati, Jane Cheung, and Chris Wong. "The last call for sustainability: An urgent growth agenda for consumer products and retail." IBM Institute for Business Value. August 2021. https://ibm. co/sustainability-consumer-products-retail
- 4 Ibid.
- 5 Ibid.
- 6 Edelman Trust Barometer 2021. Edelman. January 2021. https://www.edelman.com/sites/g/files/ aatuss191/files/2021-01/2021-edelman-trustbarometer.pdf
- 7 Gupta, Sachin, Sashank Yaragudipati, Jane Cheung, and Chris Wong. "The last call for sustainability: An urgent growth agenda for consumer products and retail." IBM Institute for Business Value. August 2021. https://ibm.co/sustainability-consumer-products-retail
- 8 Henderson, Rebecca. Reimagining Capitalism in a World on Fire. PublicAffairs. 2020.
- 9 Balta, Wayne, Manish Chawla, Jacob Karl Dencik, and Spencer Lin. "Sustainability as a transformation catalyst: Trailblazers turn aspiration into action." IBM Institute for Business Value. January 2022. https://ibm.co/sustainability-transformation
- 10 Ibid.
- 11 IBM Institute for Business Value Human Insights Global Survey of 14,526 adults. July 2021. Unpublished analysis.
- 12 Ibid.

- 13 Ibid.
- 14 IBM Institute for Business Value Trust and Transparency Consumer Survey of >14,500 consumers. November 2018. Previously unpublished data.
- 15 Goehring, Brian, Francesca Rossi, and David Zaharchuk. "Advancing AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020. https://ibm.co/ai-ethics
- 16 Guszcza, Jim and Ann Skeet. "How businesses can create an ethical culture in the age of tech." World Economic Forum. January 2020. https://www.weforum.org/agenda/2020/01/ how-businesses-can-create-an-ethical-culture-in-the-age-of-tech/
- 17 Balta, Wayne, Manish Chawla, Jacob Karl Dencik, and Spencer Lin. "Sustainability as a transformation catalyst: Trailblazers turn aspiration into action." IBM Institute for Business Value. January 2022. https://ibm.co/sustainability-transformation
- 18 Ibid.
- 19 Goehring, Brian, Francesca Rossi, and David Zaharchuk. "Advancing AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020. https://ibm.co/ai-ethics
- 20 Minevich, Mark and Francesca Rossi. "Why you should hire a chief AI ethics officer." World Economic Forum. September 2021. https://www.weforum.org/ agenda/2021/09/artificial-intelligence-ethicsnew-jobs/
- 21 Interview with Chun Schiros and Daniel Stahl, Regions Bank. November 2021; Misra, Minav. "Regions Bank: A methodology to develop high quality and trustworthy AI." IBM Blogs. May 24, 2021. https://www.ibm.com/ blogs/client-voices/regions-bank-develops-high-qualitytrustworthy-ai/

- 22 Interview with Rob Reich, Associate Director, Institute for Human-Centered Artificial Intelligence, Stanford University. October 2021.
- 23 Goehring, Brian, Francesca Rossi, and David
  Zaharchuk. "Advancing AI ethics beyond
  compliance: From principles to practice." IBM
  Institute for Business Value. April 2020. Previously
  unpublished data. https://ibm.co/ai-ethics
- 24 Ibid.
- 25 Ibid.
- 26 Ibid.
- 27 Interview with John C. Havens, Director, Emerging Technologies & Strategic Development, IEEE Standards Association. October 2021. Abdur-Rahman, Pavel. "AI & Ethics: The Grand Challenge for Our Generation." IBM Blogs—Canada. May 4, 2021. https://www.ibm.com/blogs/ibm-canada/ 2021/05/ai-ethics-the-grand-challenge-for-ourgeneration/
- 28 Edelman Trust Barometer 2021. Edelman. January 2021. https://www.edelman.com/sites/g/files/aatuss191/files/2021-01/2021-edelman-trust-barometer.pdf
- 29 Goehring, Brian and Anthony Marshall. "The business value of AI." IBM Institute for Business Value. November 2020. https://ibm.co/ai-value-pandemic
- 30 "Responsible Use of Technology: The Microsoft case study." World Economic Forum. February 25, 2021. https://www3.weforum.org/docs/WEF\_ Responsible\_Use\_of\_Technology\_2021.pdf
- 31 "Responsible Use of Technology: The IBM case study." World Economic Forum. September 2021. https://www3.weforum.org/docs/WEF\_
  Responsible\_Use\_of\_Technology\_The\_IBM\_Case\_
  Study\_2021.pdf

- 32 "Responsible Use of Technology: The Microsoft case study." World Economic Forum. February 25, 2021. https://www3.weforum.org/docs/WEF\_Responsible\_Use\_of\_Technology\_2021.pdf
- 33 Interview with Stephen Cave, Director of the Leverhulme Centre for the Future of Intelligence and a Senior Research Associate in the Faculty of Philosophy at the University of Cambridge. February 2021.
- 34 Interview with Johan Maas, CIO, and Johan Suringa, Netherland Enterprise Agency Rijksdienst voor Ondernemend Nederland (RVO). November 2021.
- 35 Interview with David Danks, Professor of Data Science & Philosophy, University of California San Diego. September 2021.
- 36 Boinodiris, Phaedra. "Getting to Trustworthy AI." VentureBeat. March 2021. https://venturebeat. com/2021/03/14/getting-to-trustworthy-ai/
- 37 Mantas, Jesus. "Empathic AI could be the next stage in human evolution—if we get it right." World Economic Forum. July 2019. https://www.weforum.org/agenda/2019/07/empathic-ai-could-be-the-next-stage-in-human-evolution-if-we-get-it-right/
- 38 Minevich, Mark and Francesca Rossi. "Why you should hire a chief AI ethics officer." World Economic Forum. September 2021. https://www.weforum.org/agenda/2021/09/artificial-intelligence-ethics-new-jobs/
- 39 Lohr, Steve. "Group backed by Top Companies Moves to Combat A.I. Bias in Hiring." New York Times. December 8, 2021. https://www.nytimes.com/2021/12/08/ technology/data-trust-alliance-ai-hiring-bias.html
- 40 "Why apply design thinking to artificial intelligence." Team Essentials for AI course. IBM. https://www.ibm.com/design/thinking/page/badges/ai

© Copyright IBM Corporation 2022

IBM Corporation New Orchard Road Armonk, NY 10504

Produced in the United States of America | April 2022

IBM, the IBM logo, ibm.com, and IBM Research are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at: ibm.com/legal/copytrade.shtml.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

This report is intended for general guidance only. It is not intended to be a substitute for detailed research or the exercise of professional judgment. IBM shall not be responsible for any loss whatsoever sustained by any organization or person who relies on this publication.

The data used in this report may be derived from third-party sources and IBM does not independently verify, validate or audit such data. The results from the use of such data are provided on an "as is" basis and IBM makes no representations or warranties, express or implied.