



Responsible AI – Maturing from theory to practice

Foreword

As awareness and scrutiny of the risks associated with AI increase, building responsible technology has become a paramount concern in organizations across all sectors. Globally, responsible AI is maturing from a “best practice” to the high-level principles and guidance necessary to drive system-level change and engender trust. Regulators are also taking notice, advocating for regulatory frameworks around AI, including increased data protections, governance and accountability measures.



In 2017 we began working on topics related to responsible AI, formally launching the Responsible AI toolkit in 2019. The suite of customizable frameworks, tools and processes was designed to help clients harness the power of AI in an ethical and responsible manner. Accompanying the launch was a survey designed to understand the key priorities, concerns and maturity of organizations attempting to deploy AI responsibly. At the time, companies were relatively new to deploying AI. Being less aware of the risks, they implemented inconsistent practices and paused at points in their journey.

Since then, entire industries have enabled better end-to-end solutions for Responsible AI¹. We surveyed over 1,000 executives across the US, UK, Japan and India to understand how views have changed — and how priorities have shifted for organizations.

We reprised some of the questions from our last [Responsible AI survey in 2018](#) and asked how the impact of recent events, such as COVID-19 and regulatory focus, has shifted priorities and heightened awareness of risks.

At a time when quick decision-making was needed, the rapid spread of COVID-19 caught governments, businesses and citizens off-guard. This prompted many businesses to accelerate AI use and innovation. Today, only 5% of our survey respondents do not currently use AI; last year, that number was 47%. The ability to operationalize AI effectively — what we call AI maturity — is key to both maintaining progress among leaders and closing the gap for companies that have yet to start their responsible AI journey.

The survey results showed where companies were grouped among the three levels of AI maturity:

25%

companies with fully embedded AI

55%

companies at the experimental stage of AI implementation

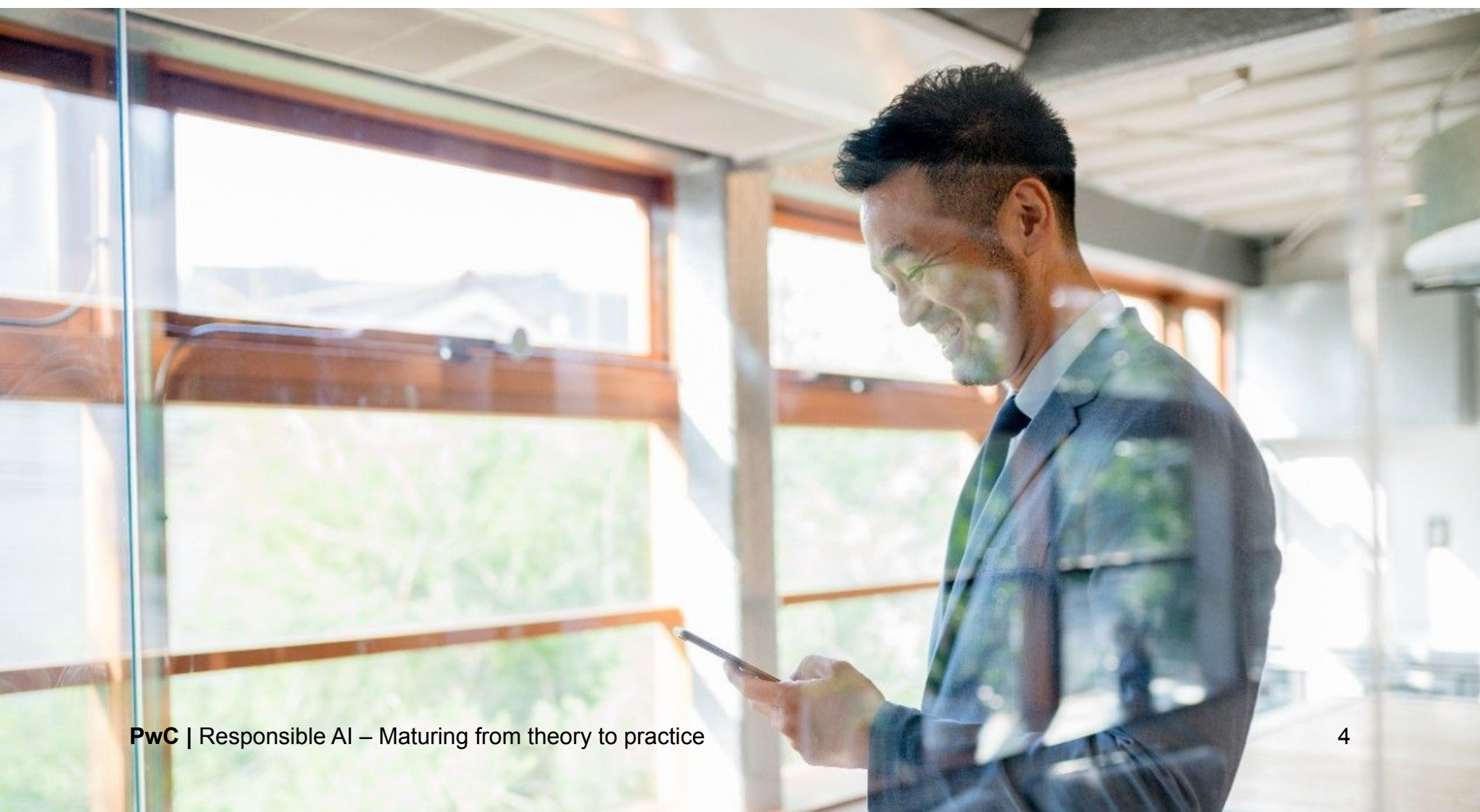
20%


companies still exploring AI without having implemented anything

¹ <https://ieeexplore.ieee.org/document/9377738>

While a quarter of companies have fully embraced AI, and more than half are experimenting, the ability to deploy at scale varies. Our research shows those with an embedded AI strategy can more reliably deploy applications at scale with increased adoption across the business. As a result of COVID-19, larger companies (>\$1bn) are significantly more likely to increase their use of AI (38%) explore new use cases for AI (39%) and train more employees to use AI (35%). The accelerated rate at which companies are adopting AI and embedding it into their organizations creates a vital need for responsible and ethical AI.

Organizations in early stages of AI adoption might be quick to look for technical fixes to potential concerns. However, those that are more mature in their adoption curve rely on a comprehensive, values-driven and tech-enabled approach to governance. In this whitepaper, you can explore the market trends surrounding ethics, risks and governance in the AI space and get a view of the rapidly changing regulatory landscape. In each section, you find key takeaways from each maturity phase along the AI adoption curve. You can also learn about some of the emerging topics and concepts that organizations across the globe can leverage in their pursuit of responsible AI.



A person with dark hair, wearing a grey button-down shirt, is seen from the back and side, sitting at a desk and looking at a computer monitor. The monitor displays a software interface with various charts and data. The background shows a blurred office environment with windows.

Updates to our responsible AI framework

Over the past two years, our views on implementing AI responsibly have evolved, based both on survey data and on our experience with clients, governments, not-for-profits and academia. For instance, there is a greater focus on helping organizations operationalize AI ethics, moving past principles to concrete practices for developers, users, and business teams. We also added more tangible practices surrounding data ethics, cybersecurity, process transparency, risk management, privacy governance, safety and sustainability.



Figure 1 – PwC’s responsible AI framework

Strategy

Data and AI ethics

Is your development, use and oversight of data and AI solutions ethical and moral?

Policy and regulation

As the regulatory landscape continues to evolve, how are you positioning your AI to meet future compliance requirements? Are you considering localized differences?

Performance and security

Bias and fairness

Is your AI fair? How are you defining that fairness?

Interpretability and explainability

Can you explain both the overall decision-making and the individual predictions generated by your AI?

Robustness

Is your AI system stable? Does it consistently meet performance requirements and behave as intended?

Privacy

How will your AI system protect and manage privacy, and how will you respond to consumers’ evolving expectations?

Safety

Is your AI safe for society? What are its potential impacts on users, and is it able to prevent unintended or harmful actions?

Security

What are the security risks and implications that should be managed to maintain integrity of algorithms and underlying data, while reducing the possibility of malicious attacks?

Control

Risk Management

Do your risk detection and mitigation practices enable the identification of emerging risks and harms across AI development and deployment?

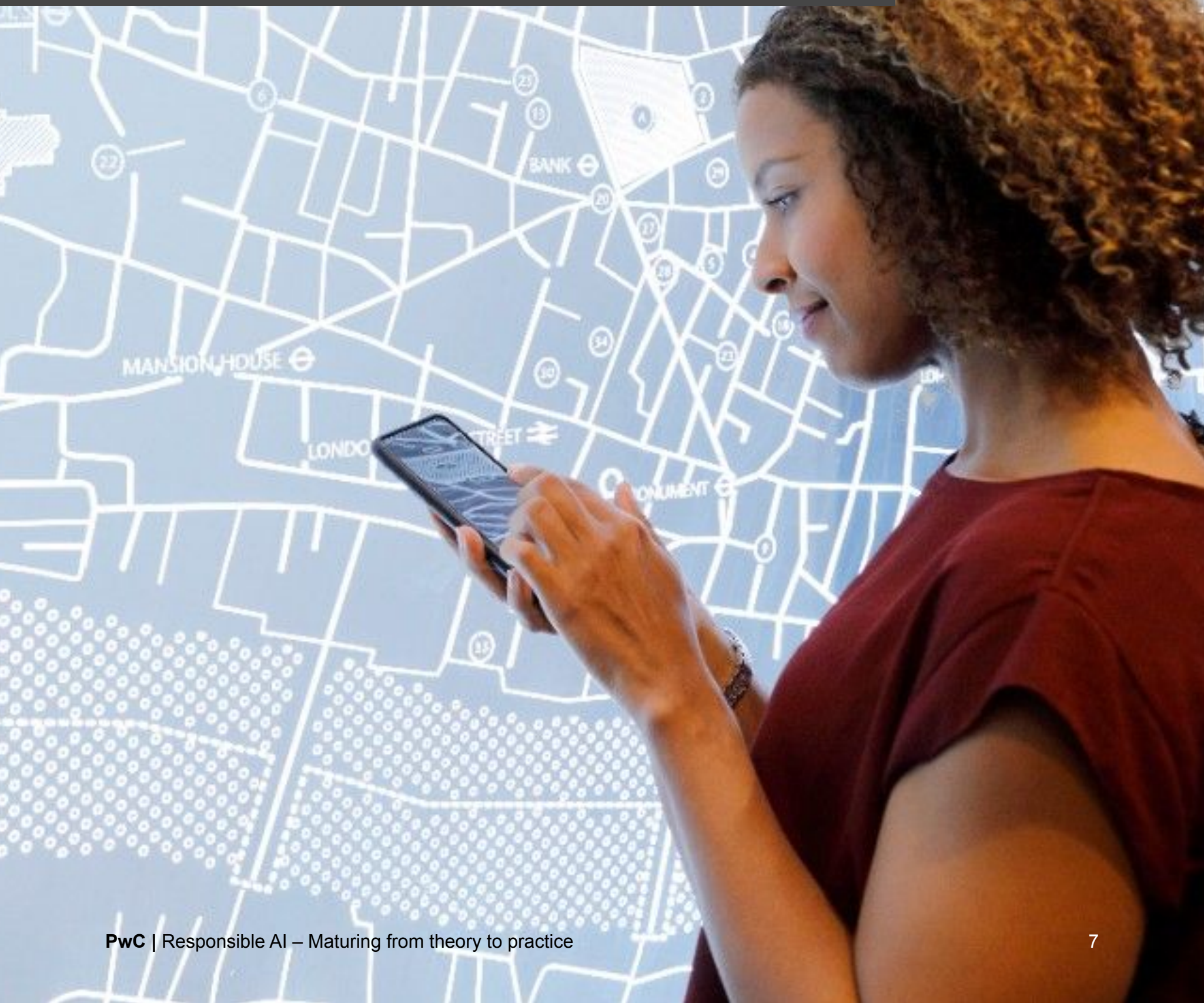
Governance

Do you have robust governance models for your AI system? Do they enable oversight with clear roles, responsibilities and requirements, as well as mechanisms for traceability and ongoing assessment?

Compliance

How are you anticipating future compliance, creating organizational policies and communicating change to stay ahead of current data protection, privacy regulations and industry standards?

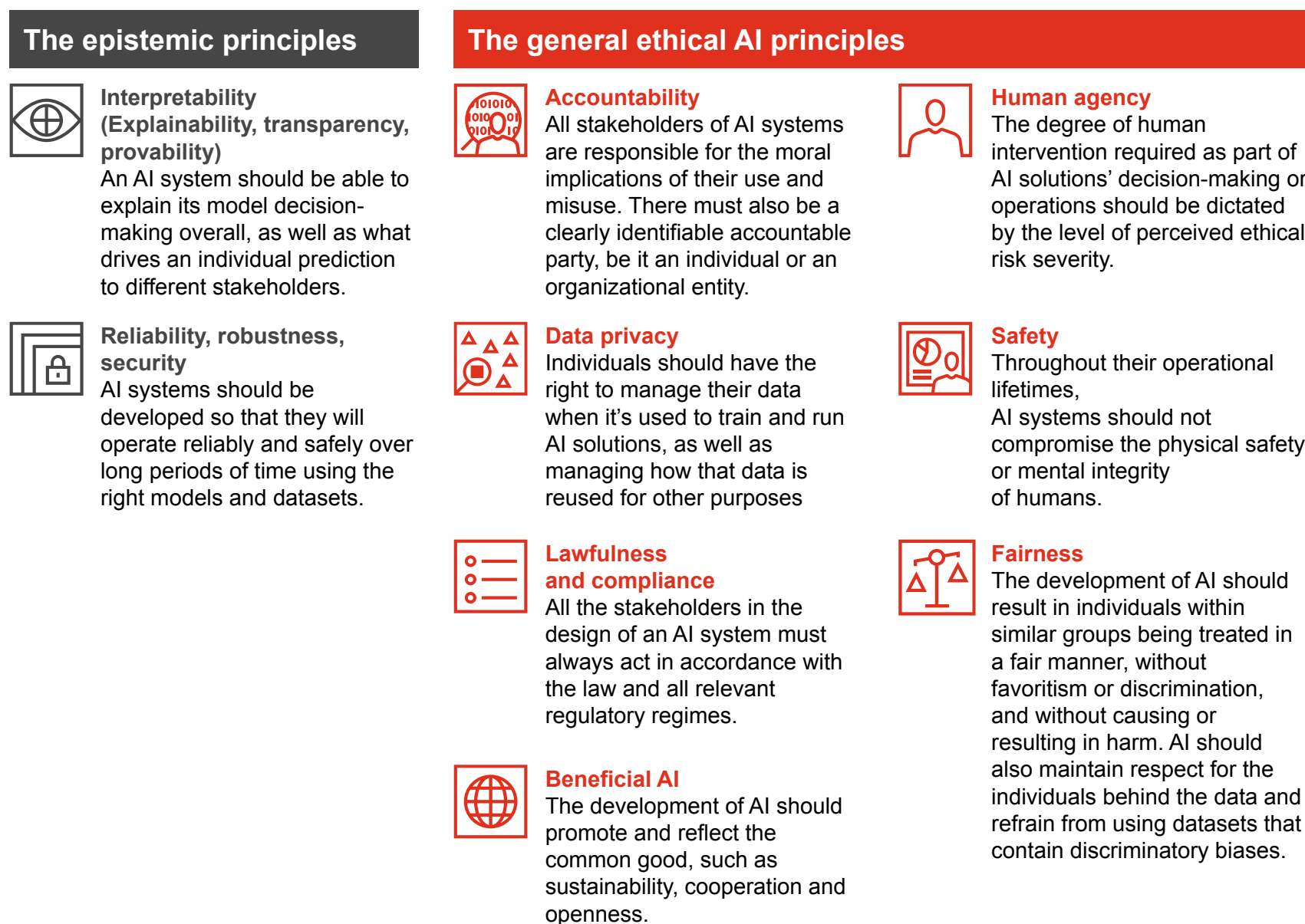
Ethics for AI is still young, but adoption is growing



AI is becoming essential across industries to help boost human productivity and decision-making, but do the benefits to the bottom line outweigh the potential impact to society? We have seen AI's disruptive potential, as well as negative consequences from its underuse, misuse and abuse.² Consumers and the media have drawn attention to biased recruitment³ and financial tools⁴, concerns around discrimination⁵ and more, raising awareness of the moral dilemmas surrounding the deployment of AI⁶. Developers, users and organizations need clear guidance and principles in order to apply AI to real-world problems responsibly and to handle all identified moral implications.

The landscape of ethical AI principles that PwC has researched is expansive and rich, but there are commonalities. We took more than 100 sets of ethical principles — amounting to 200 in total — and consolidated them into nine core ethical AI principles.

Figure 2 – Ethical AI principles



²<https://doi.org/10.1007/s11023-018-9482-5>

³<https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias>

⁴<https://hbr.org/2020/11/ai-can-make-bank-loans-more-fair>

⁵<https://ico.org.uk/about-the-ico/news-and-events/ai-blog-human-bias-and-discrimination-in-ai-systems/>

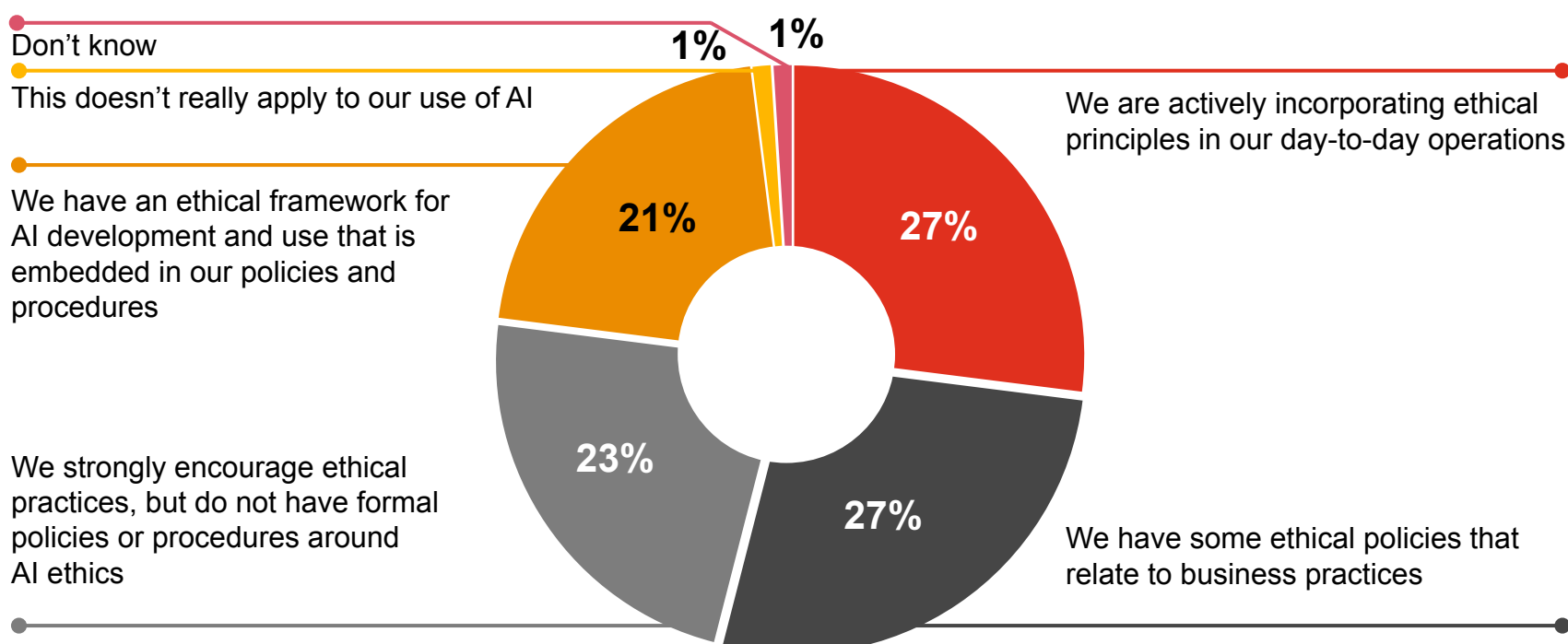
⁶<https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/>

⁷<https://www.weforum.org/agenda/2021/06/ethical-principles-for-ai/>

While releasing principles is a great first step, those principles need to be applied. Even with high expectations around ethical principles, there are few consistent approaches to put them into practice. Still, our survey shows that organizations are trying. Over half of the companies surveyed have some formal policies or principles to address ethical issues that arise from using AI, with nearly a quarter having some guidance and policies in place. This trend is consistent across countries we surveyed. One in 5 companies has an ethical framework in place for AI development and use. Encouragingly, those that are fully embracing AI in the organization are almost twice as likely to have a formal ethical framework in place (41% versus 21%).



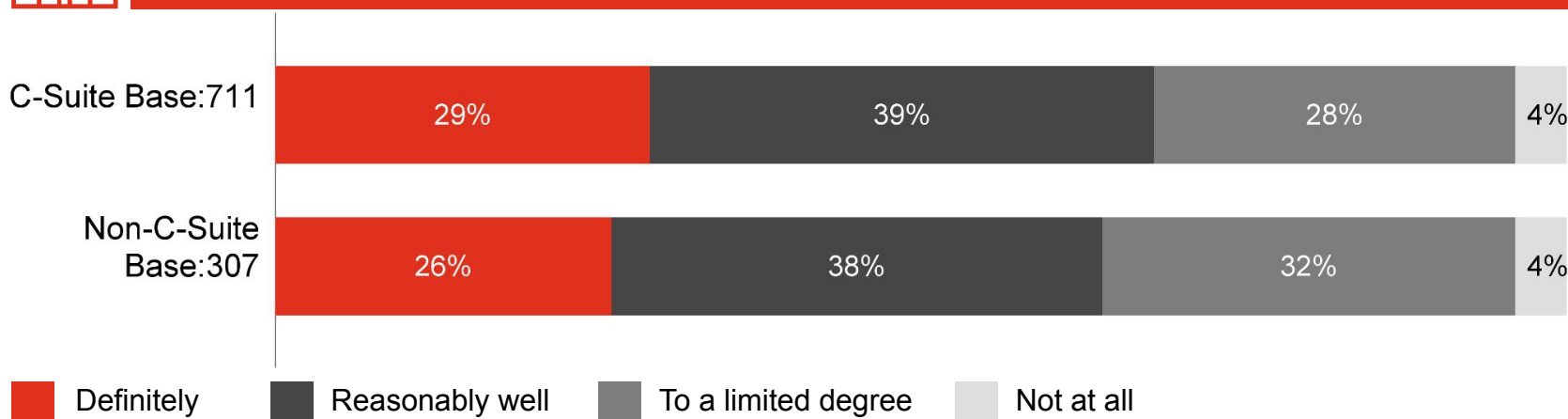
Figure 3 – State of operationalizing ethical principles



During the investment stage, executive and technical leadership support is critical for operationalizing ethics in AI in a proactive and sustainable way. 69% of CEOs and 64% of IT professionals are fairly confident about ethically driven AI investments. A majority of both C-suite (55%) and non-C-suite (47%) executives are confident that the ethics of AI aligns with organizational value, and that percentage is much higher (81% and 71% respectively) in organizations that have fully embraced AI.



Figure 4 – Support for operationalization of ethical principle by job type



Perhaps the greatest barrier to operationalizing ethics in AI is an inconsistent and linear approach. Often the initiatives launched — from AI codes of conduct to ethical boards or frameworks — are considered in isolation, which limits their ability to work effectively. While codes of conduct (63%) and impact assessments (52%) are popular tools with executives across different-size organizations, providing ethical training, using ethical boards and other means of interventions vary significantly according to organizational size and AI adoption maturity. In fact, large companies with high AI maturity use are significantly more likely to have an ethical board (60%), carry out impact assessments (62%) and provide ethical training (47%), revealing these to be the leading practices required to enable responsible AI.

But where do we go from here? Building on the promising developments we have seen across sectors and organizations of all sizes, we can acknowledge that operationalizing ethics requires resource commitment and incentives in order to deliver responsible outcomes with AI. The operationalizing of adherence to ethics can also have a massive contribution to the acceleration of AI adoption and to return on investment (ROI).

 **Figure 5 – Challenges to AI adoption (* ethical principles challenges)**

	Total
Impact of COVID on the organization is slowing AI investment	1st
Concerns with the reliability of AI applications performance over time *	2nd
Inadequate technology infrastructure to support cloud-based AI applications	3rd
Lack of the right AI technical and management talents	4th
Lack of data or poor-quality data to use in AI	5th
Concerns with data privacy *	6th
Lack of trust that an AI investment will deliver the expected returns	7th
Lack of coordination needed to make AI successful	8th
Takes too long to demonstrate the value of AI	9th
Lack of sufficient AI budget	10th
Legal concerns about our responsibility if there is an AI failure *	11th
Lack of understanding of how AI models/applications make decisions	12th
Lack of appropriate AI governance structures and frameworks *	13th
Algorithmic bias or other ethical concerns *	14th
Lack of sufficient data management policies	15th
Lack of AI-specific controls	16th











Takeaways

- While AI ethics relates to the ethical vision surrounding the use, development and objectives of AI systems, responsible AI is the multidisciplinary domain needed to translate this vision into practical guidance.
- Ethical AI frameworks should be consistent with international human rights law⁸ to support not only moral and legal accountability, but also the development of “human-centric” AI for the “common good.”
- The alternative to a piecemeal AI ethics approach (focused on individual initiatives like codes of conduct, ethical boards, ethical training and impact assessments) is a systematic overview that considers a variety of ethical AI interventions. (See figure 6.)



Figure 6 – Ecosystem of ethical AI interventions

Initiative	Description
 Value statement	Having a strong ethical vision for AI, driven by the C-suite, represents the foundation for a fair, transparent, beneficial, safe, robust outcome with AI.
 Principles and codes of conduct	The ethical principles defined by organizational values should be translated in organizational policies, codes of conduct and frameworks to allow for operationalizing those principles.
 External ethics boards	Ethical boards are part of the ethical decision-making through which ethical issues can be escalated, tensions can be managed and precedents can be set.
 Culture of ethics	Cultures are at the heart of this change and where ethical skills, knowledge and behavior should be recognized, rewarded and appreciated. Proper incentives and rewards schemes should be in place to stimulate ethical behavior.
 Education and training	Formal ethical training programs and curricula should be embraced, along with other activities that will educate individuals about ethical thinking, analysis and reasoning. These include community practice, events, book clubs, team debates and hackathons.
 Reporting/ advice channels	Having appropriate means and ways for employees to receive advice regarding ethical dilemmas or to report breaches around AI and data can help identify potential ethical issues and solve them before they escalate.
 Product development and design	Ethical decision-making and actions should be operationalized at product level, with development process ethically aligned, ethical pit stops, and check and balances embedded at every step of the process to allow for the translation of principles into norms and the norms into design and governance requirements.
 Periodic assessments	Periodic audits are necessary to assess the performance of AI in terms of fairness, safety and reliability, and where relevant areas comply with internal and external standards.

⁸<https://tech.humanrights.gov.au/>

A person is working at a computer workstation. The main monitor displays a code editor with syntax-highlighted code on a dark background. A hand is holding a blue pen, pointing towards the screen. Another monitor is visible in the background, also displaying code. The scene is lit with warm, soft light, suggesting an office or home workspace.

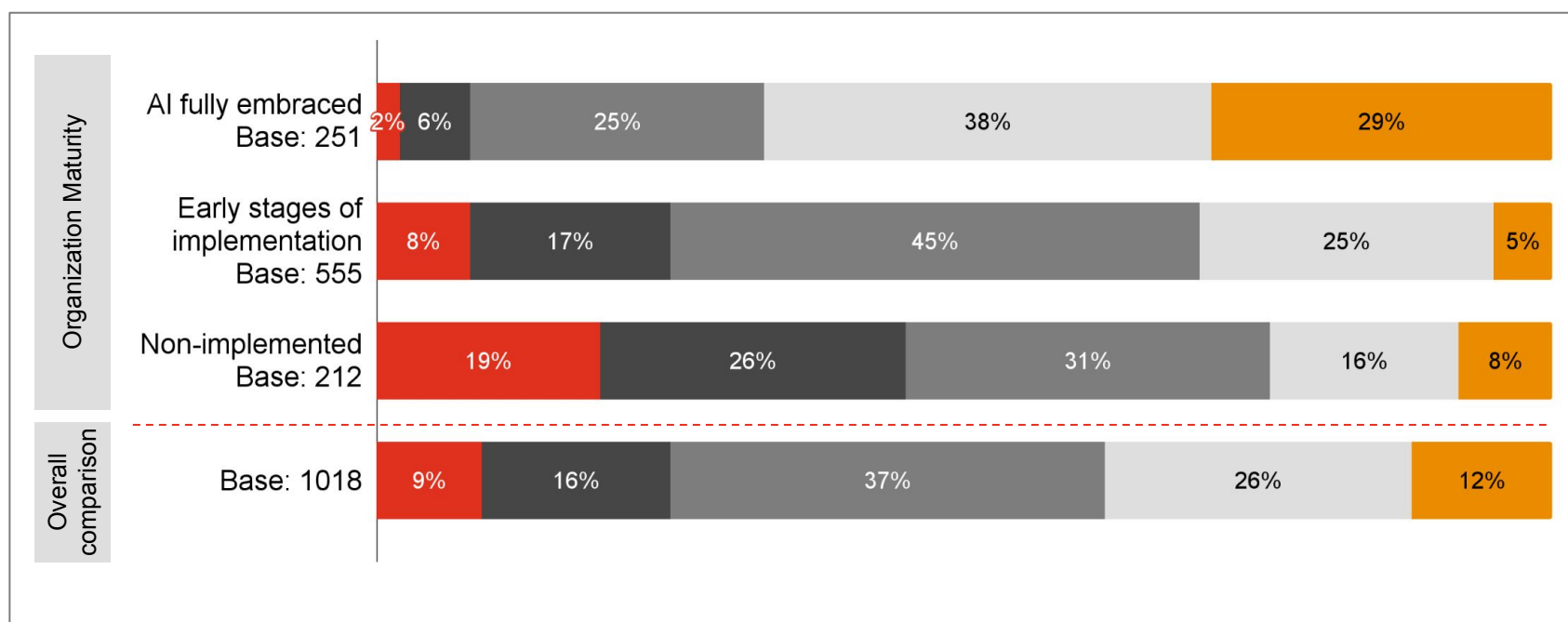
The risks of AI are a priority for businesses

Appreciation for AI ethics is growing in parallel with awareness of AI risks. This makes sense as many AI risks are in some sense ethical risks. Our survey indicated organizations are increasingly prioritizing AI risk identification and accountability, mainly by adopting an enterprise approach to AI risk mitigation. In fact, more than a third of companies (37%) have in place — and communicate — a strategy and policies to tackle AI risk. This is a stark increase from 2019’s 18%. Another quarter of companies have an enterprise approach to AI risk that is not only communicated, but also standardized.



Figure 7 – Focus on AI risk identification by maturity level

Currently, how are AI risks identified in your organization?



- No formal approach to AI risk evaluation
- Scattered silo based approach to AI risk
- Strategy and policies to tackle AI risk are in place and communicated
- An enterprise approach to AI risk has been developed, communicated and standardised
- AI risk management and internal controls are fully embedded and automated

The taxonomy of AI risks includes those at the application level — like performance, control and security risks — as well as those at the broader ecosystem level — including enterprise, societal and economic risks. Some of the more prominent and visible risks are those impacting performance, which include poor accuracy and the presence of errors stemming from poor data quality, bias, overfitting or inadequate testing procedures⁹.

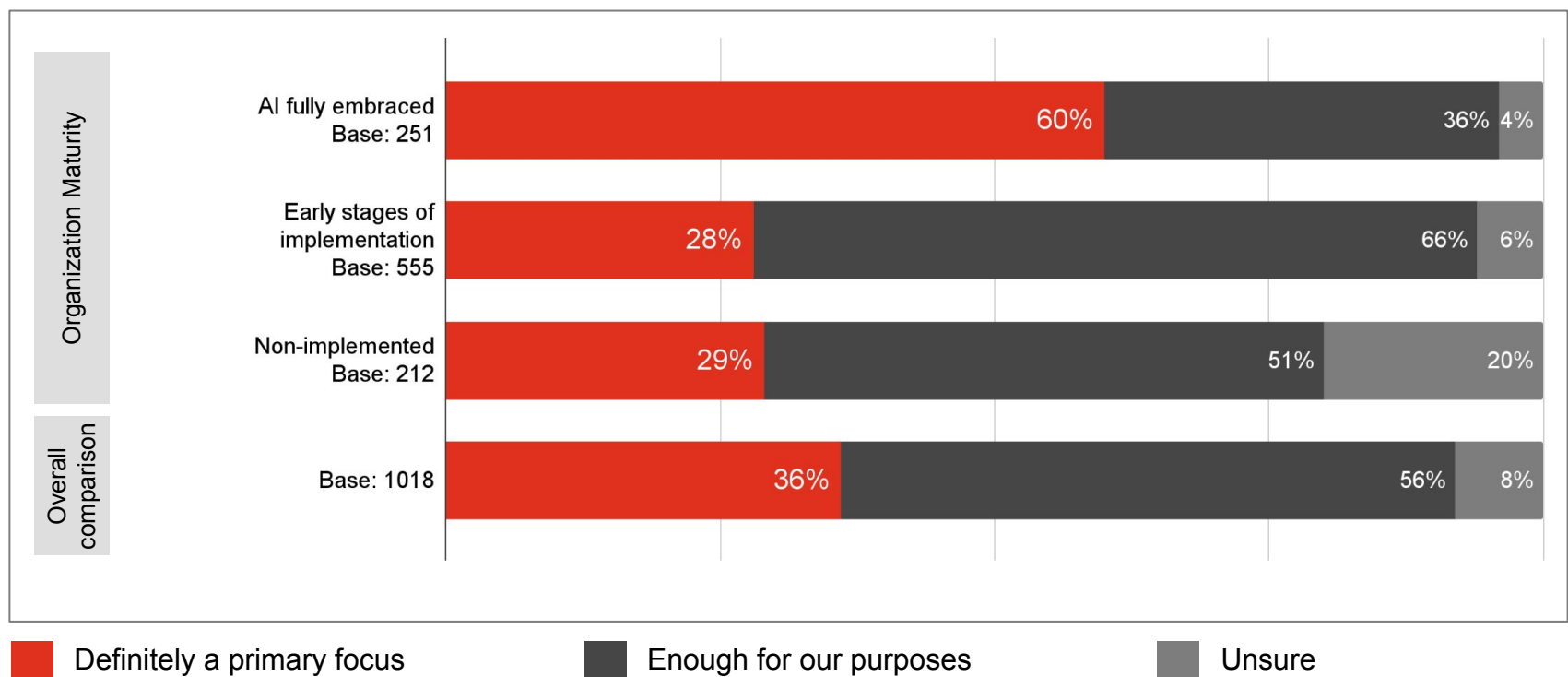
⁹<https://www.techuk.org/resource/with-great-power-comes-great-responsibility-the-importance-of-proactive-ai-risk-management.html>

Algorithmic bias is a primary concern for many organizations, partially due to emerging regulatory frameworks, media and consumer attention on discrimination, and consequent high reputational risk, as well as a desire to “do the right thing.” Respondents to our survey agree: 36% say algorithmic bias is a primary risk focus area, and 56% of respondents believe they can address bias risks adequately. Maturing companies embrace algorithmic bias as a primary focus (nearly 60% of AI leaders) as they gain more tangible experience in developing AI and awareness of issues around AI risks. Mature organizations also place a higher importance on the Fairness principle (5th place versus 8th place for less mature firms).



Figure 8 – Focus on bias by maturity level

Over the last 12 months, has your organization taken specific account of algorithmic bias (systems creating unfair outcomes, such as privileging one group over another, including gender, race or ethnicity, etc.) for your AI solutions?



Our study also shows that sensitivity to the topic of bias varies across countries. The share of companies that declare algorithmic bias as a primary focus is higher in India (48%) and the US (39%) where the public debate is significant and the population has a complex racial and ethnic composition.

Bias is often cited together with opaqueness, an inability to understand how the system makes a decision, as the main concerns that hinder AI adoption and responsible use. Opacity emerges from the growing complexity of algorithms and techniques, which can result in poor understanding of how applications function, their most important characteristics and their causal effects. Lack of transparency and poor communication can create risks by misleading users and jeopardizing the trust in an organization.

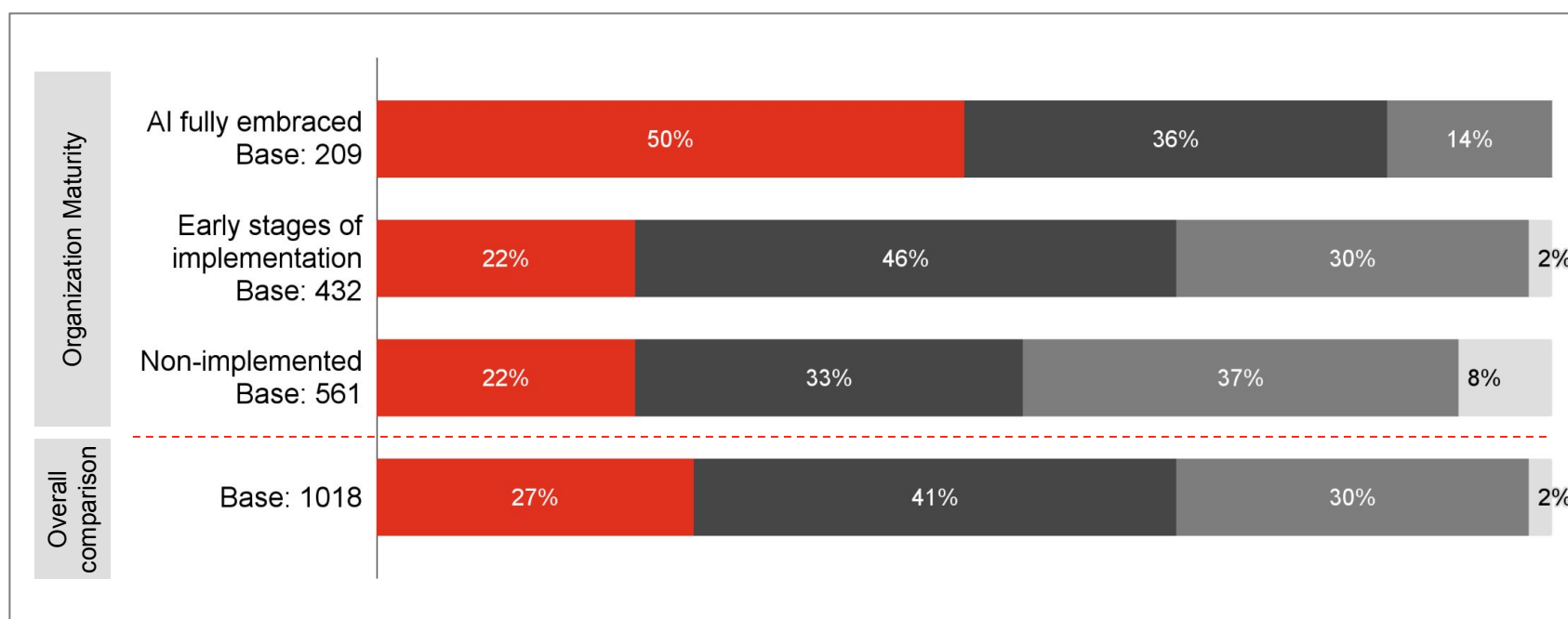
Tackling these challenges is a complex process entailing technical tools, robust processes, and, at times, a trade-off between performance and explainability. In this case, our survey respondents agree: Only 27% claimed they definitely had the ability to explain or justify the decision made by a model, while 41% could explain reasonably well and 30% could explain to a limited degree.

The picture changes if we analyse the differences across the maturity levels of AI adoption: When AI is fully embraced, half of the companies can definitely explain their decision, while only a fifth of less-mature organizations can do the same. Across the sectors, companies in Health industries are well above average (36% of respondents state “definitely” versus 27% of all respondents) on the explainability maturity, likely due to the high severity of potential harm connected to wrong decisions in healthcare.



Figure 9 – Focus on explainability by maturity level

If asked, would you be able to explain or justify a decision made by an AI in your business unit?



Once an AI system has been developed and tested, and its performance is found acceptable for its purpose, the real challenge is to enable performance stability over time. Our survey respondents are very clear about that: “Reliability, robustness and security” ranks first or second place between the ethical principles across all sectors and all maturity levels, and it is identified as the second inhibitor to AI adoption — about 10% of the respondents selected it.

This brings up another risk category for AI systems: security risks. While some of these risks may be comparable with those of other IT systems, AI increases both their probability and severity. Some security risks can emerge directly as a result of the AI techniques used. For example, adversarial attacks on machine learning models can maliciously induce an AI system to misclassify or incorrectly predict something — like convincing a computer it is seeing a toaster when it is actually looking at a banana¹⁰. And data poisoning can maliciously compromise data sources used for training so that an AI system begins to act unexpectedly¹¹.

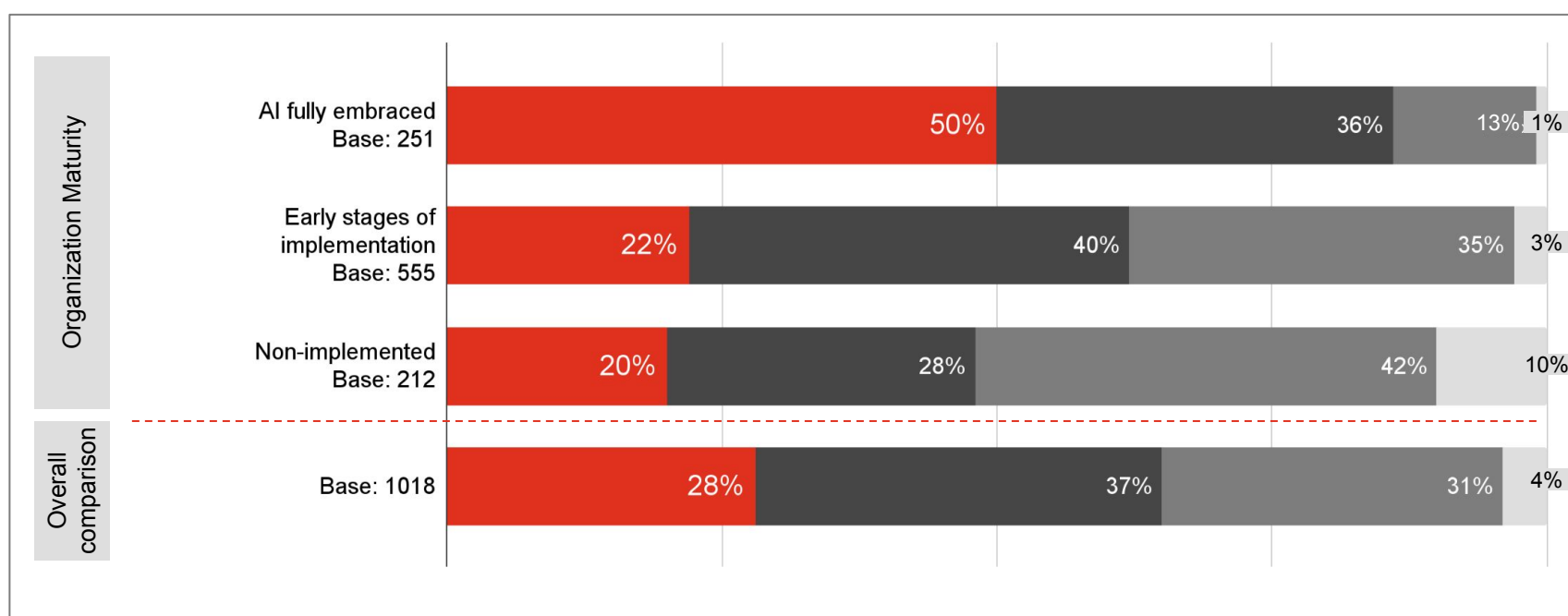
Moreover, don't underestimate the risks posed by the use of open source software. Although many open source systems can be reliable tools, there may be issues with stability and reliability over time because these tools are designed so that almost anyone can contribute to them and make changes. Consequently, the software may look different from one moment to the next, which impacts how organizations assess the risks of these tools.

While survey respondents indicated that safety was a primary concern, only more mature organizations reported that they had the ability to detect and then shut down a malfunctioning system. In fact, half of the AI leaders of mature organizations feel very confident about this capability, while only around 20% of AI leaders of less-mature organizations feel the same way.



Figure 10 – Focus on safety by maturity level

Currently, how confident are you in your organization's ability to detect and then shut down a malfunctioning AI system in a timely manner, i.e., before any serious problems are caused?

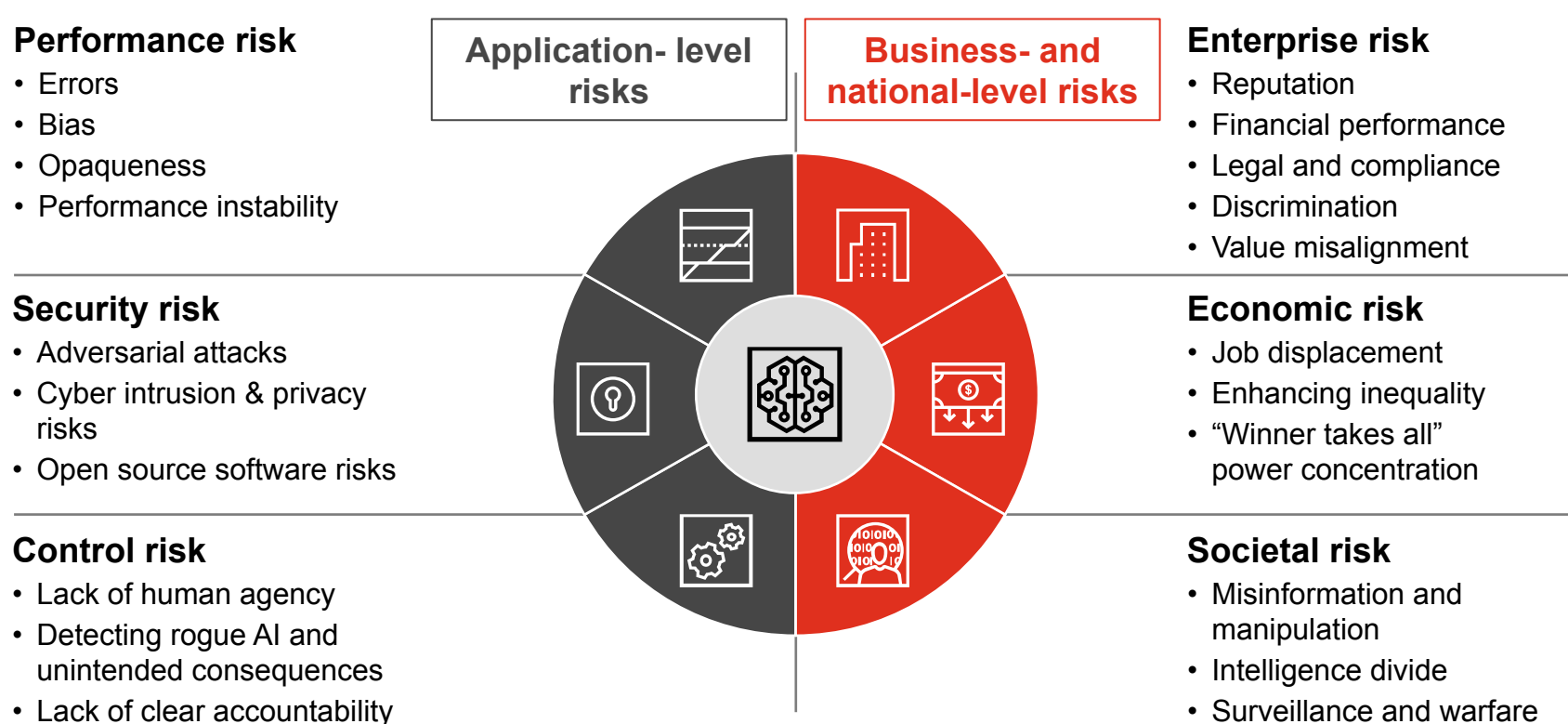


¹⁰<https://www.vox.com/future-perfect/2019/4/8/18297410/ai-tesla-self-driving-cars-adversarial-machine-learning>

¹¹<https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>

Performance is not the only risk area that should be top of mind for AI. For example, our survey shows that AI leaders recognise the skills mismatch in the current workforce as a primary concern: The lack of the right AI technical and management talent always ranks in the top five concerns across all territories, sectors and different maturity levels. The skills gap carries significant risks for the workforce in terms of job displacement and disqualifying tasks, but it also increases the risk of poor quality AI systems and difficulties managing third parties.

 **Figure 11 – AI risk categories**



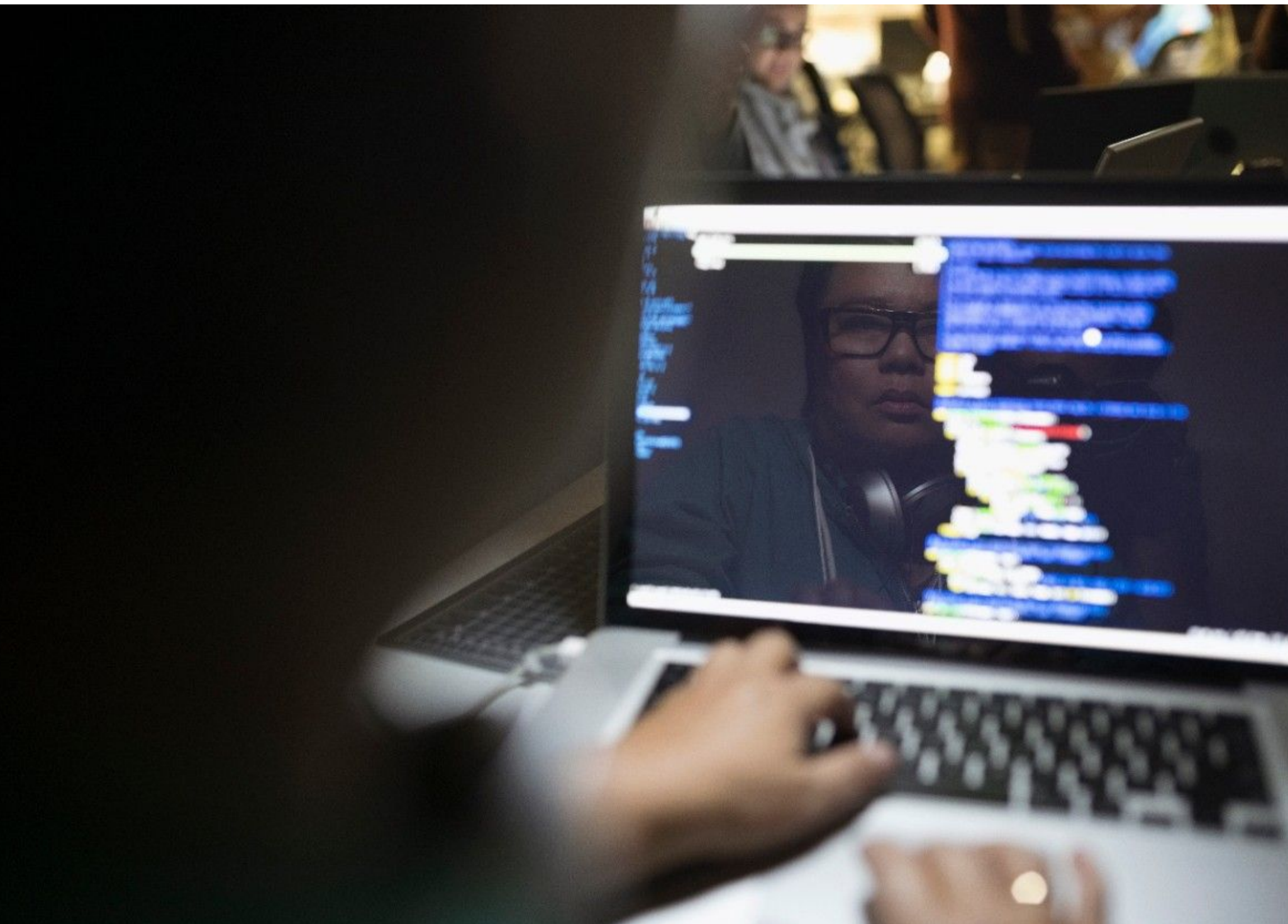
Broader organizational-level risks stem from AI’s potential for automation and economic advantages, as well as from the application-level risks described earlier. These include enterprise risks, such as reputational and financial loss, risks for non-compliance with legal requirements, discrimination, misalignment with corporate and societal values, and the management of third parties and partners.


Ultimately, the vulnerability of certain scenarios and the priority of risks vary by company, industry and type of use case. For example, risks stemming from poor reliability and stability over time are highly important in the health sector, while the public sector should be especially concerned about human rights and meeting higher compliance standards. Organizations should take an enterprise-level approach to identify AI risks, and then manage and mitigate them in a flexible way while considering the real context of the application.



Takeaways

- The specific risk an AI system may pose is directly related to its application context. For example, ask what data is used, what types of decisions are made and by whom, and which AI technique has been adopted.
- Organizations need an enterprise approach and risk management procedures to identify, evaluate, mitigate and monitor the AI risks over time.
- Make the whole organization aware of AI risks, how they can occur and how to mitigate them.
- Special attention should be given to bias and interpretability risks.





Organizational AI governance is coming of age

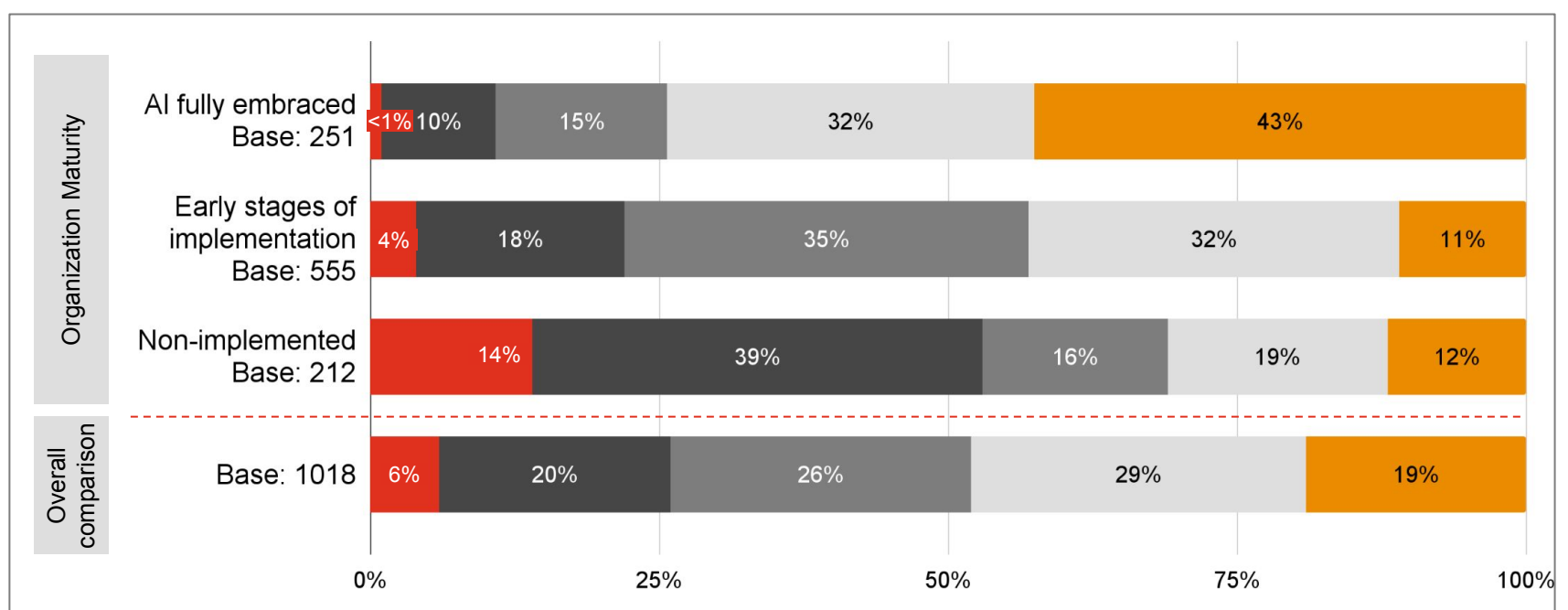
Keeping pace with shifting regulatory landscapes, effectively mitigating risks, and enacting policies that align with contextualized data and AI ethics for an organization require robust governance and accountability. Practically, this means that organizations not only need to identify tools and structures to oversee the development of AI, they also need to define accountable parties for AI use, development and oversight. Increasingly, organizations are turning to technical solutions to mitigate bias, improve explainability, monitor for robustness over time and more. It is important to note that these tools are at different levels of maturity, and their use may not fully satisfy the needs outlined by ethical principles. A holistic approach to governance uses process, policies and standards, and holistic governance that is tech-enabled rather than simply tech-first.

Our survey results showed AI risk identification and accountability is still in its infancy. Only 19% of participating companies have a formal and documented process that gets reported to all stakeholders; 29% of companies have a formal process only when there is a specific event; and the rest have only an informal process or no clearly defined process at all.



Figure 12 – AI accountability by maturity level

Currently, how do you think AI accountability is identified in your organization?



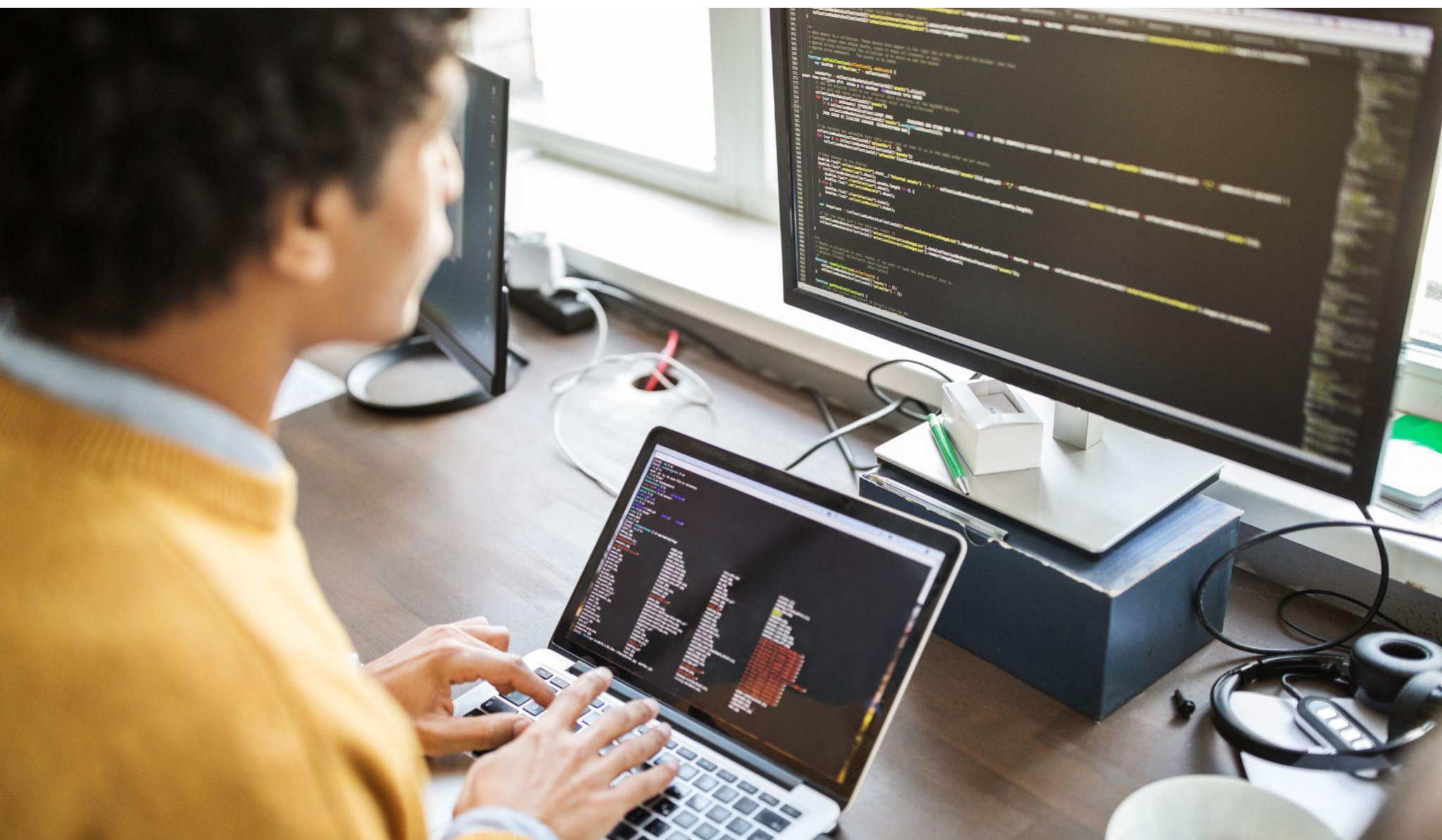
- We do not have any clearly-defined process for identifying AI accountability
- The developer or the person who approved the application or model for release is ultimately accountable
- We have an informal review process that is triggered if there is an incident
- We have a formal, documented review process that is triggered if an incident occurs
- We are completely transparent about the boundaries of accountability for all stakeholders and gather or publish documented evidence to support this transparency

Effective risk mitigation and governance require an organization-wide approach to AI and data governance. This approach must be end to end, as well as far-reaching across all three lines of defense, where ownership and accountability are clearly articulated.



End-to-end governance

Holistic governance begins with the strategy of an organization, which should include the desired uses and expectations for data, analytics and AI. At this level, organizations need to define their priorities. The planning stage is where organizations stand up the programs for model development and data use, followed by the ecosystem stage, which sources the technology and personnel required to achieve the targets established in the strategy phase. Governance should be proportional, meaning it should be tied to the context of the application itself¹² so it doesn't impose overly burdensome tasks on the development teams or stifle innovation. Several factors may inform the governance requirements, including the risk of the system itself, the privacy of the data used, the novelty of the system and need for new governance mechanisms.



¹²https://www.pwc.com/jp/en/knowledge/thought_leadership/comprehensive-ai-governance-needed-now.html



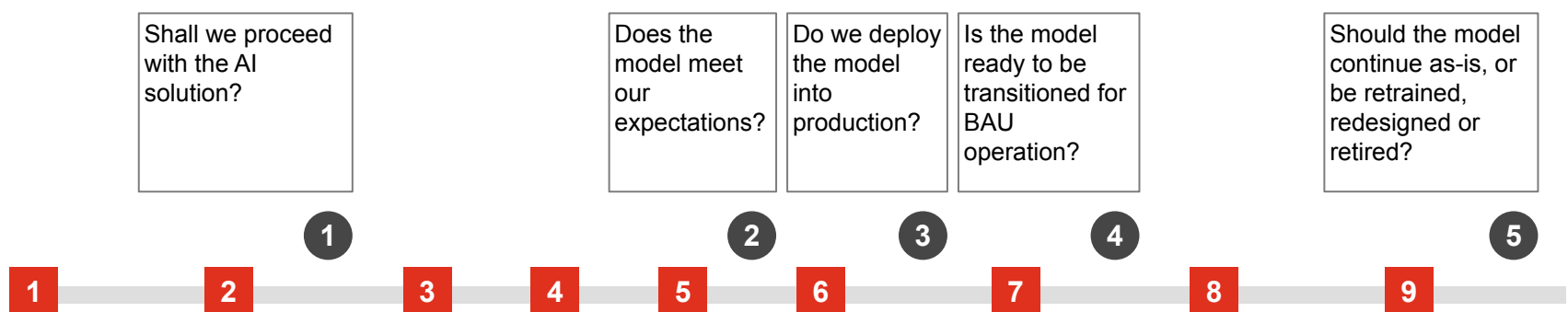
Figure 13 – The development life cycle for analytics and AI

Individual AI systems are developed in the iterative nine-step model development and deployment processes. In these stages, data scientists and developers must translate business needs and priorities into well-scoped models and software processes. Data must be obtained, transformed and manipulated as needed for the application. A model is iteratively built, trained and tested, until an optimal solution is determined. This solution is independently validated against user expectations and existing processes before being formally deployed. Then it is continually monitored for efficacy.

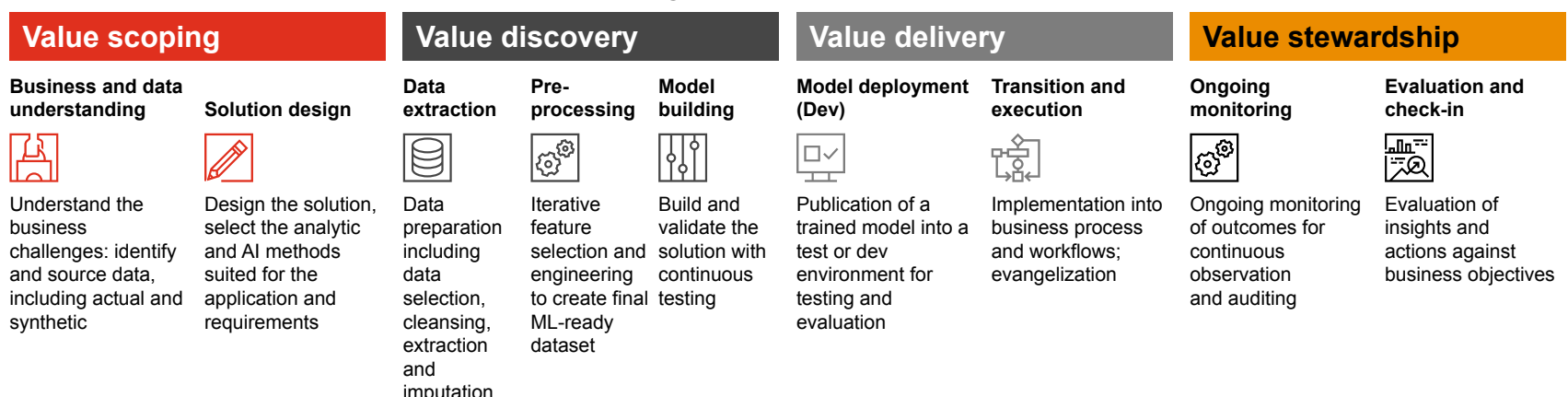
Technical solutions from software and cloud companies target these nine steps for their governance solutions. At each stage gate, the development and business teams (sometimes referred to as the First Line) work with leadership and quality assurance reviewers (the Second Line) to obtain sign off. It is not a given that a system will pass all stage gates. In fact, to do so requires testing, documentation and alignment on expectations. The development process is accompanied by policies, standards and procedures that are complied with. Importantly, this life cycle does not constitute all of the governance process for AI. That process is larger and starts with organizational strategy.¹³ A third line, Internal Audit, evaluates the effectiveness of controls.

These five stage gates are designed to engage the three lines at different points in the development process, making the conscious decision at each point to progress the application based on concrete requirements.

Stage gates



Nine-step model development life cycle



¹³ <https://towardsdatascience.com/top-down-and-end-to-end-governance-for-the-responsible-use-of-ai-c67f360c64ba>



Algorithmic Impact Assessment (AIA) makes an entrance

One mechanism that can be leveraged across the five stage gates — which has been increasingly called for by the [academic community](#) and has been proposed in regulations — is the Algorithmic Impact Assessment (AIA). While there are different interpretations of what AIAs look like in practice, they are ultimately intended to accompany the development of a system, to make critical decisions about the design and limitations of a given system, and to act as a source of documentation for others (like consumers) who would ultimately benefit from or be impacted by the system in deployment. Some organizations are creating new templates for an AIA, while others are building off a Data Protection Impact Assessment, which is already required under the EU's General Data Protection (GDPR) for risky processing of data.



Governance tools pick up steam

Other areas experiencing significant growth are the governance tools released specifically for the AI and data science community. These take a few forms:

Built-in capabilities for monitoring within cloud or deployment platforms

On-demand assessment tools oriented toward black box or white box evaluation

Documentation automation and workflows for development.

These capabilities will be useful for effective governance, but they do not provide complete governance on their own. The use of these tools needs to be dictated by the organization so there is consistency of evaluation, monitoring and controls, while making sure they serve the same objectives. For example, different teams using different bias assessment tools with different definitions of fairness may develop tools and systems that these teams believe are meeting the objective of fairness, but they may actually be in conflict with one another.



Takeaways

- Define organizational guidelines and standards for governance that business units can leverage.
- Incorporate the three lines of defense structure into AI development and escalate sensitive uses to cross-functional teams for review.
- Apply consistent documentation templates and criteria to improve transparency.
- Use governance tools as enablers for decision-making.

A woman with short brown hair, wearing a dark green shirt, is seated at a white table in a modern office setting. She is gesturing with her right hand while speaking to a man with glasses and a beard, who is wearing a dark patterned shirt. The man is looking towards her. In the foreground, the back of a person's head and shoulders is visible, suggesting a meeting or discussion. The background shows large windows and office equipment, creating a professional atmosphere.

A rapidly shifting policy and regulatory landscape



Public policy is increasingly playing an important role in addressing the challenges of balancing benefits with risks, and clarifying how and where AI should be used in certain public contexts and how it should be regulated¹³. Our study participants ranked lawful and compliant application of AI as their third priority, following reliability and data privacy — moving up from fourth place.

Some of this policy is driven by increased advocacy toward addressing identified risks, such as discrimination and inequality. At both the national and supranational level over the past few years, rich, robust AI public policy initiatives have emerged, engaging stakeholders across the public and private sectors, academia, research, regulators, think tanks, advocacy and standards, among others. At the national level, many countries have issued National AI and data strategies to boosting innovation, research and development, offering business and consumer protection, training and reskilling¹⁴. At the supranational level, emerging policies for AI are designed to provide robust guidelines and recommendations to govern technology and to align with ethical principles and human rights. These efforts include those driven by institutions ranging from the European Commission on Trustworthy AI, the Organization for Economic Co-operation and Development (OECD) and consortia like the Global Partnership on AI.

The most important roles public policy can play are to be a source for robust, focused and agile regulation, and to signal leading practices to organizations. Because regulations are often slow to launch, regulators engage in public policy efforts to inform soft-law and guidance. Once policy is tested in the market, it may inform a more robust regulatory framework. Many companies have expressed a desire for more concrete regulatory requirements: For example, major technology companies have backed away from facial recognition in the US until regulation is passed. Many policymaking bodies are investing significantly in upskilling, research, onboarding industry experts, formal industry collaboration and advocating for sandboxed development.

¹³ <https://digitaltechitp.nz/2021/04/07/how-do-we-ensure-the-responsible-use-of-ai-by-governments/>

¹⁴ <https://www.pwc.lu/en/advisory/digital-tech-impact/technology/gaining-national-competitive-advantage-through-ai.html#:~:text=At%20PwC%20C%20we%20have%20developed.being%20made%20by%20different%20countries.>



By forming strong partnerships between public policy bodies and industry, some policymakers are trying to complement existing self-governance approaches with regulatory frameworks, considering the use cases and risks involved¹⁵. For example, the UK and European Union are outlining approaches in which AI ethics principles will guide the overall design, development and deployment of AI in a country, complementing a graded risk-based approach that varies by use case and by sector.

All countries are different, so a one-size-fits-all approach to policy would not work well. As such, AI-related regulations are in various stages of maturity across the world economies. These regulations are often built on common themes of data privacy and protection, accountability and innovation in AI.

Providing data protection and privacy is a leading theme in AI-related policies and regulations that are picking up steam. While the landmark GDPR was passed by the EU several years back, other territories have launched similar initiatives. For example, bills at the state level across the United States focus on data privacy for AI and automated decisioning systems. Japan is defining a legal framework around data collection and usage as shown by its modification of the Act on Protection of Personal Information and the introduction of the Act on Anonymously Processed Medical Information to Contribute to Research and Development in the Medical Field.

Another emerging regulatory theme is accountability, in which a mechanism growing in popularity is a risk-based approach to governance. The European Union's GDPR has acted as a template for the recently released AI Act¹⁶. This act also introduced conformity assessments, common in the software space, for quality assessment. In the US, proposals like the Algorithmic Accountability Act also reference assessments as a way to provide accountability.

To pursue innovation while protecting against the risks, regulations focused on specific use cases are emerging, as well as the sponsorship of standards to guide development.

¹⁵ <https://www.weforum.org/whitepapers/ai-governance-a-holistic-approach-to-implement-ethics-into-ai>

¹⁶ <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>



The recent Request for Information by the five financial services regulators in the US highlights potential upcoming financial services-specific guidance. Besides bills, the US has issued the Executive Order on Maintaining American Leadership in Artificial Intelligence, and has devised several draft technical standards around bias, accountability, explainability and more for AI systems with the National Institute of Standards and Technology (NIST).

Japan's attempt at removing hurdles in AI adoption and balancing R&D acceleration is evidenced by the amendments of the Copyright Act, which allows the business model for selling learning data sets used for machine learning, and the Road Transport Vehicle Act, which paves the way for autonomous driving. Singapore builds on this trend with the Road Traffic Autonomous Motor Vehicles Rules (to regulate trials of autonomous motor vehicles and push the development of automated vehicle technology) and the Protection from Online Falsehoods and Manipulation Act (to advocate punishment for fake news and deep fakes).

Countries like Australia, which do not yet have any specific laws regulating AI or algorithmic decision-making, are advancing a range of related use-case-specific laws and legal concepts for autonomous vehicles and autonomous weapons systems. These countries may also be advancing national AI strategies and agendas, like Australia's AI Action Plan, Digital Economy Strategy, AI Ethics Framework and AI Technology Roadmap, and AI Standards Roadmap.

A commonality across these regulatory efforts is the push to prioritize ethics and develop trust with the subjects of data collection who, invariably, are also potential consumers of AI. Some countries consider fundamental human rights, social empowerment and Sustainable Development Goals as foundational to proposed regulation, as evident in recent proposals from India¹⁷ and Japan¹⁸. It appears these efforts intend to prevent the throttling of innovation in the enforcement of these regulations.

¹⁷<https://niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf>

¹⁸<https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>



Proposed EU legislation on AI¹⁶

The European Union has proposed a first-of-its-kind regulatory framework for the ethical use of AI (the AI Act). The EU proposal divides AI systems according to categories of risk and imposes different restrictions for each category.

1

The proposal outright bans any kind of AI activity that could seriously impact individuals. These AI technologies pose unacceptable risks according to EU standards.

2

The proposal has identified several categories of high-risk AI and has laid separate rules for their usage: critical infrastructure, education/vocation, safety components, employment and worker management, and biometric identification among others. Articles 6 through 11 of the regulation establish classifications of high-risk AI, their compliance and the specifics of their risk management system, data and governance rules, technical documentation and record-keeping. Human oversight is compulsory. The regulations also apply to providers of high-risk AI apart from developers. Providers are responsible for establishing and maintaining quality management systems, drawing up technical documentation and maintaining compliance with all regulations. These risk categories are hotly debated at the moment.

3

The framework does not impose heavy restrictions on AI technologies that pose limited or minimal risks.

Policymakers will need to consider a number of key issues as they debate the European Commission's proposed law to regulate AI. Most agree on the goal of strengthening European competitiveness in the global economy, but views diverge on how to achieve that without putting European businesses at a disadvantage. For example, some policymakers want to see requirements for algorithmic explainability and transparency, plus ex-ante risk assessments, while others are more willing to embrace soft-law solutions such as self-regulation. In reaching a consensus on how to achieve their aim of strengthening European competitiveness in the global economy, lawmakers will need to decide how to strike the right balance between protecting consumers and encouraging innovation.

Regulatory activity around AI systems has only emerged in the past few years, which means policymakers have minimal precedent to reference when they defend the bill. While the EU may be the first to introduce a legal framework for AI systems, it's likely that others will follow.

¹⁶<https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>



Takeaways

- Policy and regulation can be accelerators (helping increase adoption by engendering trust by managing for moral implications and risks), as well as inhibitors (slowing innovation).
- Regulations built on the foundation of these policies not only ensures safe, robust and ethical use of AI, but also provides clarity about its development. This may bolster confidence and encourage R&D investment in AI by reducing uncertainty. Therefore, regulators are considering legal hurdles that do not impede innovation, but instead build trust for all involved stakeholders.
- Politics can play an important role in countries' attempts to increase their competitive edge and raise their eminence in the AI field both globally and locally. In some cases, there is incentive for governments to frame human-rights-focused, data-driven policies, since that indicates appreciation of the fundamental rights of citizens.
- Businesses are taking an increasingly important role in helping frame AI-related policies and in collaborating with regulators. This collaboration, including with civil society, is needed to develop regulations that are coherent and holistic and that create competitive advantages for future generations.
- Public policy framing and implementations are a win-win situation for governments and companies alike. Regulators utilize them as a testing ground before hard regulations are put in place. These policies also act as guidance for the things to come for executives who plan on building or utilizing AI in both their strategic and day-to-day operations.
- Algorithmic decision-making is highly contextual, and new legal protections will be needed to address the challenges that will emerge with time.

Emerging topics in responsible AI





As consumer expectations for trust and transparency grow, comprehensive **data ethics** frameworks are needed to explicitly embed values into the entire data supply chain. This will build on the data privacy and protection compliance processes that start maturing in organizations as a result of data privacy legislation, like the European Union’s GDPR and the California Consumer Protection Act (CCPA). In fact, these legislative acts provided the first step toward adoption of ethical data practices, as data privacy was selected as one of the more important ethical principles by our survey respondents. With the increase in data use across organizations, especially for AI systems, the need for an ethical approach to data management is critical. Data ethics might be the most robust approach toward achieving ethical AI, by confirming that the right ethical principles are considered in the context of data supply chain, and by building a robust, responsible foundation for future AI applications and uses.

In the absence of regulatory frameworks, assessing the quality of AI systems’ output against set standards emerges as one of the more accessible ways to govern AI responsibly — even though few standards are defined. The field of **AI assurance** is described as “governance mechanisms for third parties to develop trust in the compliance and risk of a system or organization.”¹⁹

Formal definitions surrounding assurance require robust, industry-agreed- upon standards. The community is advancing other approaches, including bias audits, certifications, accreditations and impact assessments in lieu of (or in anticipation of) standards. An international ecosystem is required to facilitate the consistency and interoperability of approaches across jurisdictions to enable a global perspective on AI governance²⁰.

There is a growing role for Second Line functions (Privacy, Compliance and Data Governance) to come together for better AI and data governance. Despite this need, there is also recognition that there likely should be a **designated owner to oversee the rollout of AI governance** and coordinate collaboration between teams. Some organizations might choose the Privacy group to own AI governance, given the increased requirements beyond just compliance that some are adopting with respect to data and data use.

¹⁹<https://cdei.blog.gov.uk/2021/04/15/the-need-for-effective-ai-assurance/>

²⁰<https://cdei.blog.gov.uk/2021/05/11/the-european-commissions-artificial-intelligence-act-highlights-the-need-for-an-effective-ai-assurance-ecosystem/>



Other organizations might choose IT ownership, given the technical nature of AI systems. Those in financial services might decide to expand the remit of Model Risk Management to oversee governance of all AI systems, not just high-risk, regulated models. Other entities are considering establishing an entirely new office for the role of AI governance. Compliance-shy organizations might also choose to bestow AI governance ownership on the First Line themselves, choosing self-regulation. This practice is not common, as the answer to “Who owns AI?” is not always the same as “Who owns AI governance and responsible AI?” The reality is that most organizations are still trying to figure out what works best for them, and there is no one-size-fits-all solution given the lack of standards currently in place. Which organizational model to adopt, which governance mechanisms and tools to use, how to connect ethical principles to practices, which external bodies to engage, and how to report are all in flux. Further, most organizations do not have centralized development teams and, consequently, do not know precisely where AI is in use across the enterprise. What is clear is that leveraging existing governance structures wherever possible can help increase adoption and decrease the inertia that sometimes accompanies processes that are viewed as burdensome and overly bureaucratic.



Contacts

Anand Rao

Global Artificial Intelligence Leader,
PwC US

Annie Veillet

Partner, One Analytics, PwC Canada

Matt Kuperholz

Partner, Chief Data Scientist, PwC Australia

Matt Labovich

Data, Analytics and Technology Leader,
PwC US

Euan Cameron

AI Leader, PwC United Kingdom

Sudipta Ghosh

Data & Analytics Leader, PwC India



To learn more, visit our Responsible AI website.
www.pwc.com/rai



© 2021 PwC. All rights reserved. PwC refers to the US member firm or one of its subsidiaries or affiliates, and may sometimes refer to the PwC network. Each member firm is a separate legal entity. Please see www.pwc.com/structure for further details.